RADIO ENGINEERING and ELECTRONIC PHYSICS

English Edition of

РАДИОТЕХНИКА И ЭЛЕКТРОНИКА

Published by the American Institute of Electrical Engineers
with the aid of a grant from the National Science Foundation

No.10 October 1961

Translated and Produced by Royer and Roger, Inc.

AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS

Established 1884

345 East Forty-Seventh Street
New York 17, N.Y.

The English edition of RADIO ENGINEERING AND ELECTRONIC PHYSICS is published by the American Institute of Electrical Engineers with the aid of a grant from the National Science Foundation. Also published under the same arrangement are the Russian electronic journals RADIO ENGINEERING and TELECOMMUNICATIONS.

RADIO ENGINEERING AND ELECTRONIC PHYSICS

(РАДИОТЕХНИКА И ЭЛЕКТРОНИКА)

Publication of the Institute of Radio Engineering and Electronic Physics,
Academy of Sciences of the USSR

Translated and Produced
by
Royer and Roger, Inc.

Translation Editor: Herbert Dern, Columbia University

Subscriptions to Radio Engineering and Electronic Physics should be sent to AIEE
Special Subscription Department
41 East 28th Street, New York 16, New York

1961 Subscription rates:

	$	£
Individuals	28.50	10
Libraries, institutes, govt. agencies	57.00	20

12 issues per annum comprising approximately 1900 pages

IDEAL RECEPTION OF OPTIMUM-CODE SIGNALS

A.I. Sharov

A method is considered for designing an ideal Kotel'nikov receiver, intended for the reception of optimal code signals.

The effect of noise on a transmitted signal is an irreversible process; therefore what the receiver accepts as a transmitted signal is in fact a combination of signal and noise. All possible oscillations of this type, which are picked up by the receiver as transmitted signals, comprise an infinite set defined in a certain domain known as the signal domain [1]. Different receivers designed for the reception of like signals are characterized by different signal domains. Using a geometrical representation, these domains have various sizes and configurations. It is obvious that it is always advisable to have such a receiver, i.e., such signal domain dimensions and configurations, as to provide the maximum probability of correct reception. A receiver that satisfies this requirement can be called an ideal receiver.

Ideal receivers, designed for the reception of different code signals, are characterized by various signal domains with various sizes and configurations, and of two domains having the same configuration, the one with the greater dimensions will provide a higher probability of correct reception. Different signal-domain configurations having the same dimensions (volume) will provide different probabilities of correct reception. In view of this, there certainly should exist some kind of "ideal" configuration of the signal domain, which for a definite type of noise and a constant signal-domain volume provides the maximum probability of correct reception. It is easy to show that where there is "white noise" type of interference the ideal configuration will be a sphere. However, it is impossible to construct a receiver which would be characterized by n-dimensional signal domains of spherical configuration with $n > 1$ when the number of signals is $N > 2$. But for a given n there exists a closest approximation of the signal-domain configuration to spherical, namely an "optimal" configuation, which is achieved by the closest packing of spheres of equal radii in the n-dimensional space and which is an n-dimensional polyhedron in which the "closely packed" sphere is inscribed.

The ideal receiver characterized by such optimal signal domains will be intended to receive signals in some optimal code, which for a specified entropy and specified signal-to-noise ratio ensures maximum probability of correct reception and, approaches in its properties the ideal Shannon code as n is increased.

Signals of limited power, i.e., confined to the sphere in the geometrical interpretation, can be subdivided into signals with volume-spherical distribution and signals with surface-spherical distribution. The construction of an optimal code in which the signals have volume-spherical distribution requires a knowledge of the closest packing of the spheres in the n-dimensional space. On the other hand, the construction of an optimal code in which the signals have surface-spherical distribution also requires a knowledge of the closest packing of the spheres, but with the condition that all the centers of the spheres lie on a sphere of radius $\sqrt{nP}$ (P is the signal power).

Signals with surface-spherical distribution have the advantage over signals with volume-spherical distribution in that they do not require a constant signal level to be maintained at the receiver location.

The closest packing for a surface-spherical distribution is presently known only for n $\leqslant$ [2]. However, it is possible to use the closest packing for a volume-spherical distribution, which is already known for n $\leqslant 10$. Discarding the inner spheres from the closest packing fo the volume-spherical distribution leads to the formation of an approximation to a closest pack ing for surface spherical distribution; this approximation improves with increasing n. A cod with surface-spherical signal distribution, constructed in the manner indicated above from ar optimal code with volume-spherical distribution, can also be considered optimal if we take in consideration the existing possibilities of effecting closest packing with a surface-spherical d tribution.*

In this article we are considering the construction of an ideal receiver of signals in pre- precisely such an optimal code with surface-spherical signal distribution.

The basic operation performed by the receiver is to identify the received signal, dependi on the location of its representative point in signal space, with one or another of the signals t can be transmitted. This identification is accomplished by dividing the signal space into a series of nonoverlapping signal domains that are set in correspondence with the code signals

The division of the signal space into nonoverlapping domains can be made by specifying boundaries of these domains.**

These boundaries are segments of hyperplanes. In this case the domain boundaries can b specified in terms of their own equations in the initial coordinate system

$$\begin{aligned} &l_1(X) + C_1 = 0, \\ &l_2(X) + C_2 = 0, \\ &\cdots\cdots\cdots \\ &\cdots\cdots\cdots \\ &l_m(X) + C_m = 0, \end{aligned} \tag{1}$$

where $l_i(X)$ is the linear form of the n-dimensional vector $X = (X_1, X_2, \ldots X_n)$. C_i is a constant.

If the point of the received signal corresponding to the vector $X = X^0$ is not located in the plane of the domain boundary, then substitution of the coordinates of this point into the equatio for the domain boundary $l_i(X) + C_i = 0$ will result in one of two inequalities $l_i(X^0) + C_i > 0$ or $l_i(X^0) + C_i < 0$, depending on what side of the plane of the domain boundary the initial poi is situated. On this basis, substitution of the coordinates of the point of the received signal i the corresponding equation for the domain boundary makes it possible to determine, from the combination of the signs of the results obtained, the domain where the point of the received si nal is located, and therefore also what code signal is transmitted.

In the simplest case this can be seen in an example wherein equally probable binary code signals are received. The ideal receiver will then be characterized by signal domains with boundary equations.

$$\begin{aligned} &X_1 = 0, \\ &X_2 = 0, \\ &\cdots\cdots \\ &\cdots\cdots \\ &X_n = 0. \end{aligned} \tag{2}$$

---

*For n = 3, with the same number of signals N = 12 there exists a better code, construc ed through the use of spherical closest packing [2]. In this case our code is not optimal in th sense stated.

**When the code signals are equally probable, there is another possible definition of the signal domains of an ideal receiver, namely the aggregate of the points located closer to the given point corresponding to the transmitted (standard) signal than to others.

To determine the position of the point of the received signal relative to one of the boundaries of the domains it is necessary to know only one coordinate of the point, one element of the code combination. Therefore, it is possible in this case to employ successive utilization of the elements of the code combination, usually called sequential or element-by-element reception of the signal. In the general case (1), to determine the position of the point of the received signal relative to the domain boundaries it is necessary to know all the coordinates of the point, all the elements of the code combination, simultaneously. Unlike the element-by-element reception of signals, this method of reception can be defined as simultaneous utilization of the elements of the code combination, as the reception of the signal as a whole.

In the reception of signals with a surface distribution, the ideal receiver will be characterized by signal domains whose boundaries have the following equations

$$\begin{aligned} &l_1(X) = 0, \\ &l_2(X) = 0, \\ &\cdots\cdots \\ &\cdots\cdots \\ &l_m(X) = 0. \end{aligned} \tag{3}$$

Unlike the case of reception of signals with binary code (2), it is necessary to receive the signal as a whole and, compared with the general case (1), the absence of the constants C_i makes it simpler to remember (fix) the equations of the boundaries of the signal domains in the receiver. As can be seen from the system (3) it is necessary to fix in the receiver m equations. But, if the signal domains are symmetrical about the initial coordinate system, then it is possible to reduce the number of domain-boundary equations that must be fixed in the receiver. This can be done by symmetrically mapping the subspaces of different quadrants of the initial system of coordinates on the subspace of one of them, say the positive quadrant. In this case we can confine ourselves to fixing only the equations of the boundaries of the domains of the positive quadrants, regarded as symmetrical mappings of the domain boundaries situated in the other quadrants.

The placement of the signal domains characterizing the ideal receiver depends on the placement of the code signals. Symmetrical placement of the domains will take place only if the code signals are symmetrically placed (if they are equally probable).

The main feature of the placement of the points of the optimal-code signals lies in the fact that they form a regular but not orthogonal point lattice, which upon suitable choice of the initial coordinate system results in a symmetrical placement of the signal points.

In ideal reception of the optimal-code signals with a surface-spherical distribution, the signal domain boundaries can be represented by the system of equations

$$\begin{aligned} &B_1^{(1)} X_1 + B_2^{(1)} X_2 + \cdots + B_n^{(1)} X_n = 0, \\ &B_1^{(2)} X_1 + B_2^{(2)} X_2 + \cdots + B_n^{(2)} X_n = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &B_1^{(m)} X_1 + B_2^{(m)} X_2 + \cdots + B_n^{(m)} X_n = 0, \end{aligned}$$

where the coefficients $B_1^{(i)}$, $B_2^{(i)}$ $\ldots$ $B_n^{(i)}$ are the coordinates of a certain n-dimensional vector perpendicular to the plane of the i-th domain boundary, and can be evaluated as the differences in the coordinates of points corresponding to code signals situated in neighboring domains having a common boundary $B_1^{(i)} X_1 + B_2^{(i)} X_2 + \ldots + B_n^{(i)} X_n = 0$. In the case of m signal-domain boundaries, it is necessary to make m substitutions of the coordinates of the point of the received signal in m domain-boundary equations, as a result of which one obtains a combination of signs consisting of m binary elements (+ or -), which determines the domain where the points of the received signal are situated.

Since the optimal-code signals and hence also the signal domains characterizing the ideal receiver are symmetrically situated, we can confine ourselves to an analysis of only the equations of domain boundaries located in the positive quadrant. The number of domain-boundary equations which must be fixed in the receiver will undoubtedly be reduced in this case. Since we fix the equations of domain boundaries in the positive quadrant, it is necessary to substitute

into these equations the coordinates not of the point of the received signal itself, but of its mapping on a subspace of the positive quadrant. Consequently, the receiving unit must perform an additional operation, namely the symmetrical mapping of the received-signal point. In addition, for a complete determination of the position of this point with respect to the signal domains it is necessary to know its position relative to the quadrants of the initial coordinate system. It is readily determined from the combination of the signs of the coordinates of the point of the received signal.

Thus, the position of the point of the received signal relative to the signal domains is determined from the combination of the signs of the coordinates of the point (position of the point relative to the quadrants of the initial coordinate system) and from the combination of the signs which is obtained by substituting the coordinates of the symmetrical mappings of the received signal point into the equation for the domain boundaries in the positive quadrant (position of the point relative to the signal domains inside the quadrant.)

In the case of a double-valued optimal code (n = 2) having six points (N = 6) with the following coordinates:

Point number	1	2	3	4	5	6
X_1	1	-1	-2	-1	$+1$	2
X_2	$\sqrt{3}$	$\sqrt{3}$	0	$-\sqrt{3}$	$-\sqrt{3}$	0

(distance between neighboring points D = 2), we have three domain-boundary equations

$$-X_1+\sqrt{3}X_2=0,$$
$$X_1=0,$$
$$X_1+\sqrt{3}X_2=0.$$

Let us assume now that the point of the received signal A has coordinates $X_1 = -1$ and $X_2 = +1$ (Figure 1). The coordinates of the symmetrical mapping of the point A on the subspace of the positive quadrant (point A') will be $X_1 = 1$ and $X_2 = 1$ (Figure 2). Further, substituting the coordinates of the point A' into the domain-boundary equation of the positive quadrant, $-X_1 + \sqrt{3X_2} = 0$, we obtain

$$-1+\sqrt{3}\simeq 0.73>0.$$

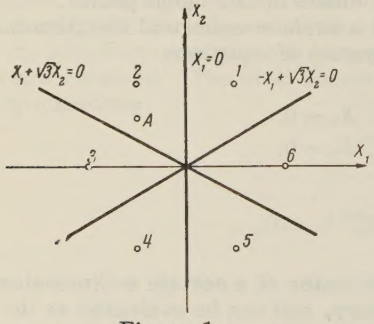


Figure 1

Then the combination of signs determining the region where the point A is located will consist of a combination of signs, defining the quadrant where the point A is located (- +), and the combination of signs defining the location of point A' in the positive quadrant (+), i.e., from -+ and +. Denoting + as 1 and - as 0, we obtain the binary combination 011, which determines the region of location of the point of the received signal, and which can be set in correspondence with the code signal.

In the case of a three-valued "optimal" code (see the first footnote), (n = 3) with twelve signal points (N = 12), we have six domain-boundary equations. In the positive quadrant we have three signal points with coordinates

Point number	1	2	3
X_1	1	2	0
X_2	1	0	2
X_3	$\sqrt{2}$	0	0

and two domain boundaries with equations

$$-X_1 + X_2 + \sqrt{2}X_3 = 0,$$
$$X_1 - X_2 + \sqrt{2}X_3 = 0.$$

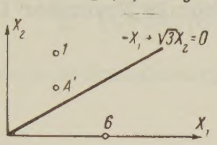


Figure 2

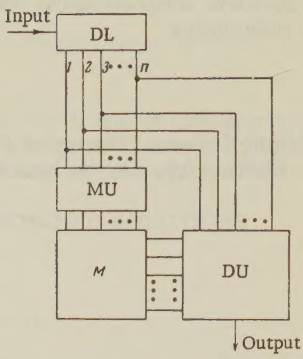


Figure 3

In the case of a four-valued optimal code (n = 4) with 24 signal points (N = 24) we have 20 domain-boundary equations. The positive quadrant will have five signal points with the following coordinates:

Point number	1	2	3	4	5	Point number	1	2	3	4	5
X_1	1	2	0	0	0	X_3	1	0	0	2	0
X_2	1	0	2	0	0	X_4	1	0	0	0	2

and four domain boundaries with equations

$$-X_1 + X_2 + X_3 + X_4 = 0,$$
$$X_1 - X_2 + X_3 + X_4 = 0,$$
$$X_1 + X_2 - X_3 + X_4 = 0,$$
$$X_1 + X_2 + X_3 - X_4 = 0.$$

In the case of a five-valued optimal code (n = 5), which already has 40 signal points (N = 40), we have 20 domain-boundary equations. In the positive quadrant we have seven signal points with the following coordinates:

Point number	1	2	3	4	5	6	7
X_1	1	2	0	0	0	1	0
X_2	1	0	2	0	0	1	0
X_3	1	0	0	2	0	0	1
X_4	1	0	0	0	2	0	1
X_5	0	0	0	0	0	$\sqrt{2}$	$\sqrt{2}$

and eleven domain boundaries.

It is seen from the foregoing examples that by using the operation of symmetrical mapping we can when $n \leqslant 5$ reduce the number of domain-boundary equations to be fixed in the receiver by two or three times, and consequently simplify the receiver itself.

A receiver operating on this principle does not contain any complicated devices. This can be seen from the block diagram of Figure 3. Here the signal as a whole is received with the aid of a delay line with taps (DL). The operation of symmetrical mapping is performed be a two-way rectifier in the mapping unit (MU). The position of the symmetrical mapping

of the received-signal point relative to the domains of the positive quadrant is determined with the aid of a matrix network (M) in which are fixed the equations of the positive-quadrant signal domain boundaries. Somewhat more complicated is the decoding unit (DU). It does, however, perform simple logical operations and can be made up of standard elements of electronic computers.

REFERENCES

1. V.A. Kotel'nikov, Theory of Potential Noise Immunity, GEI, 1957.
2. A.A. Kharkevich, On the best code, Elektrosvyaz', 1956, 2, 65.

Received by editor 29 November 1960

OSCILLATIONS IN AN OSCILLATOR WITH DELAY

P.A. Perepelyatnik

The over-all behavior of the oscillator is determined with the delay varied over a wide range. It is shown that, in the case of large delay time, the oscillations cannot in principle be stable and self-modulation set in. The hypothesis is advanced that in the case of considerable delays the oscillator turns into a noise generator.

To analyze the stability of the periodic solutions of second-order nonlinear equations with an argument subject to delay, to which this system reduces, a qualitative method is used based on the principle of violation of the alternation of the roots (the Mikhaylov criterion).

1. DERIVATION OF "ABBREVIATED" EQUATIONS

Figure 1 shows a diagram of the oscillator investigated. It is merely a model, since the deductions obtained can be extended to include an oscillator in which the principle of delayed feedback is used to produce the oscillations. To simplify the analysis we assume the following: 1) the delay element draws no current; 2) there is no reaction of the load on the amplifier; 3) the grid currents are disregarded.

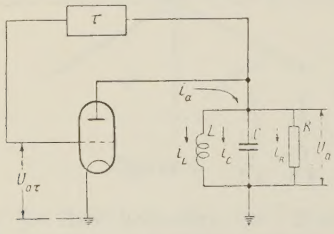


Figure 1. Diagram of oscillator with delay.

Using the symbols shown in Figure 1, we obtain the following equations

$$\frac{d^2U_a}{dt^2} + 2\delta\frac{dU_a}{dt} + \omega_0^2 U_a = -2\delta R\frac{di_a}{dt}. \qquad (1)$$

Here $2\delta = 1/RC$, $\omega_0 = 1/LC$.
We put

$$i_a = S_1 U_{a\tau} - S_3 U_{\tau a}^3, \quad U_a = Y\cos\psi,$$
$$U_{a\tau} = Y_\tau\cos\psi_\tau = Y_\tau\cos(\psi - \Delta\psi) \qquad (2)$$

Substituting (2) into (1), we carry out harmonic linearization, discard the terms containning the fundamental harmonic, perform certain trigonometric transformations, and equate the coefficients of cos ψ and sin ψ on the left and right sides of (1), after which we obtain the following two abbreviated equations:

$$2\dot{Y}\dot{\psi} + Y\ddot{\psi} + 2\delta Y\dot{\psi} - 2\delta K_{\xi\tau}\dot{Y}_\tau \sin\Delta\psi + 2\delta K_{Y\tau}Y_\tau\dot{\psi}_\tau \cos\Delta\psi = 0,$$
$$\ddot{Y} - Y\dot{\psi}^2 + 2\delta\dot{Y} + \omega_0^2 Y + 2\delta K_{\xi\tau}\dot{Y}_\tau \cos\Delta\psi + 2\delta K_{Y\tau}Y_\tau\dot{\psi}_\tau \sin\Delta\psi = 0. \quad (3)$$

Here $K_{Y\tau} = R\left(S_1 - \frac{3}{4}S_3Y_\tau^2\right)$; $K_{\xi\tau} = R\left(S_1 - \frac{9}{4}S_3Y_\tau^2\right)$; K_Y is the transfer function due to the nonlinearity, with respect to the fundamental of the high-frequency oscillation; K_ξ is the transfer function due to the nonlinearity, with respect to perturbations of the envelope.

2. INVESTIGATION OF STATIONARY PERIODIC SOLUTIONS

In the case of a stationary mode we have

$$\ddot{Y} = \dot{Y} = 0, \qquad Y = Y_\tau = Y_0,$$
$$\psi = 0, \qquad \psi = \psi_\tau = \omega, \qquad \Delta\psi = \omega\tau. \quad (4)$$

In the general case, $\omega \neq \omega_0$.

Substituting (4) into (3) we obtain

$$1 + K_Y \cos\omega\tau = 0,$$
$$\omega_0^2 - \omega^2 + 2\delta K_Y\omega \sin\omega\tau = 0. \quad (5)$$

The equations (5) can be written in a different form:

$$\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0} = \frac{1}{Q}\sqrt{K_Y^2 - 1}, \quad (6a)$$
$$\operatorname{tg}\omega\tau = Q\left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0}\right), \quad (6b)$$

Equation (6a) can also be rewritten as

$$W(Y_0, \omega) = \frac{K_Y}{\sqrt{1 + Q^2\left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0}\right)^2}} = 1. \quad (6c)$$

Here $W(Y_0, \omega)$ is the transfer function of the fundamental in the closed feedback loop.

From (6) we can obtain several important properties of the system.

1. Equation (6a) determines the frequencies ω_{max} and ω_{min} that can be excited in system. For this purpose we put $Y_0 = 0$; then $K_{Ymax} = RS_1$. Knowing ω_0, Q, and K_{Ymax}, we can readily obtain from (6a) the values of ω_{max} and ω_{min}. The bandwidth $\Delta\omega_{max} = \omega_{max} - \omega_{min}$ will, henceforth, be called the transmission band of the system. It is determined not only by the parameters of the resonant circuit, but also by the parameters of the amplifier. Using (6a), we can determine the transmission band directly from the expression

$$\Delta\omega_{max} = \frac{\omega_0}{Q}\sqrt{K_{Y\,max}^2 - 1}. \quad (7)$$

2. From (6c) we can also readily obtain the frequency dependence of the amplitude of the oscillations. For this purpose we first determine

$$K_Y = \sqrt{1 + Q^2\left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0}\right)^2}. \quad (8)$$

Determining K_Y from (5), we can readily plot the sought dependence.

3. As the delay time is varied from zero upward, there are at first no oscillations in the system; oscillations set in only at certain $\tau_{cr\,1}$ (we shall call this the first critical value of the

delay time). As τ is increased, the regions with and without oscillations will alternate. If the conditions for the existence of periodic oscillations can be satisfied in the passband of the system for two frequencies simultaneously, that is, when conditions (6) are satisfied, the alternation of regions with increasing τ will cease. This value of τ will be called the second critical value of the delay time, τ_{cr2}. For any τ greater than τ_{cr2} there will always exist oscillations in the system.

If the transmission band of the system is sufficiently large, the alternation of the regions may not take place, since the value of τ_{cr2} then falls in the interval of τ defining the first region of oscillations.

The values of τ_{cr1} and τ_{cr2}, as well as the intervals of τ which determine the extent of the regions where oscillations exist, can be determined from the expression

$$\tau = \frac{1}{\omega}\left[\operatorname{arc\,tg}\, Q\left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0}\right) + (2n+1)\pi\right], \quad n = 0, 1, 2, \ldots, \tag{9}$$

which is obtained from (6b).

Equation (9) is used to plot Figure 2, which explains the foregoing. The intervals between the starts of the τ curves corresponding to the boundaries of regions with different numbers of oscillations tend with increasing τ to a value $\Delta\tau = 2\pi/\omega_{max}$. The intersection between the horizontal line corresponding to a specific value of the delay time and the curve $\tau(\omega)$ determines the number of possible oscillation frequencies. At sufficiently large τ, the intervals between

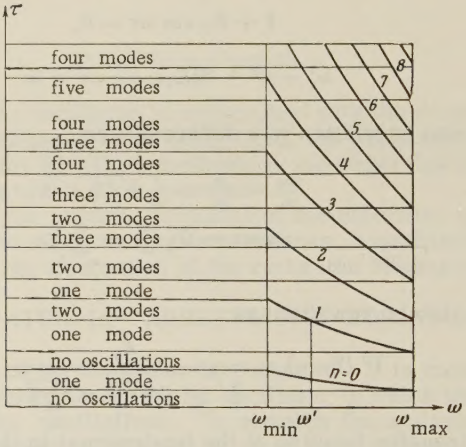


Figure 2. Curves showing the values of the delay time as a function of the frequency.

frequencies will be given, according to (9), by the expression $\Delta\omega = 2\pi/\tau$.

Putting n = 0 and $\omega = \omega_{max}$, we can obtain from (9) the first critical value of the delay time

$$\tau_{cr1} = \frac{1}{\omega_{max}}\left[\operatorname{arc\,tg}\, Q\left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0}\right) + \pi\right]. \tag{10}$$

Using (9) we can write an analytic expression for the second critical value of the delay time. Indeed the following equations hold true for τ_{cr2}:

$$\tau_{cr2} = \frac{1}{\omega_{max}}\left[\operatorname{arc\,tg}\, Q\left(\frac{\omega_0}{\omega_{max}} - \frac{\omega_{max}}{\omega_0}\right) + (2m+1)\pi\right], \tag{11a}$$

$$\tau_{cr2} = \frac{1}{\omega'}\left[\operatorname{arc\,tg}\, Q\left(\frac{\omega_0}{\omega'} - \frac{\omega'}{\omega_0}\right) + (2m-1)\pi\right]. \tag{11b}$$

Here m is the value of n for which the conditions where oscillations can exist at two frequencies begin to be satisfied. The meaning of ω' is clear from an examination of Figure 2. By suitable transformation of (11a) and (11b) we obtain

$$\tau_{cr2} = \frac{2\pi - (\varphi_2 - \varphi_1)}{\omega_{max} - \omega'}. \qquad (12)$$

Here

$$\varphi_1 = \text{arc tg } Q\left(\frac{\omega_0}{\omega_{max}} - \frac{\omega_{max}}{\omega_0}\right),$$

$$\varphi_2 = \text{arc tg } Q\left(\frac{\omega_0}{\omega'} - \frac{\omega'}{\omega_0}\right).$$

To estimate τ_{cr2} we can put $\omega' = \omega_{min}$. If we know m, then τ_{cr2} is most easily determined from (11a). It is easily seen that as the delay time increases the number of possible oscillations increases rapidly, and occupies the frequency spectrum lying in the passband of the system.

Let us examine as an example

$$K_Y = R\,(10^{-3} - 10^{-5}\,Y_0^2), \quad Q = 50, \quad C = 200\,\mu\mu f, \quad \omega_0 = 2\pi \cdot 10^6 \frac{1}{sec}.$$

From this we find R = 40 kilohms. Using (6a), we obtain

$$\omega_{max} = 2\pi \cdot 1.48 \cdot 10^6 \frac{1}{sec}, \qquad \omega_{min} = 2\pi \cdot 0.68 \cdot 10^6 \frac{1}{sec}.$$

The passband is $\Delta\omega_{max} = 2\pi \cdot 0.8 \cdot 10^6$ sec^{-1}. From expression (9) we obtain for n = 0 and $\omega = \omega_{max}$ the first critical value of the delay time, $\tau_{cr1} = 0.173$ microsecond.

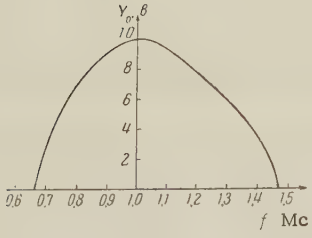


Figure 3. Dependence of the amplitude of oscillation on the frequency for the case Q = 50, C = 200 micromicrofarads, $\omega_0 = 6.28 \cdot 10^6$ sec^{-1}, $K_Y = R(10^{-3} - 10^{-5}Y_0^2)$.

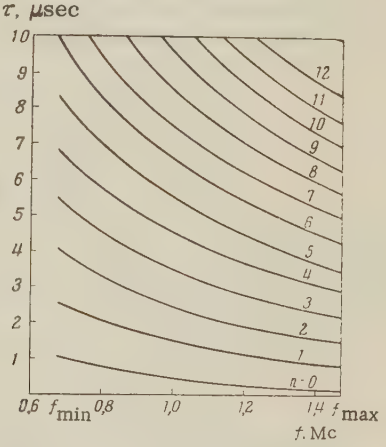


Figure 4. Curves showing the values of the delay time as a function of the frequency for the case Q = 50, C = 200 micromicrofarads, $\omega_0 = 6.28 \cdot 10^6$ sec^{-1}; $K_Y = R(10^{-3} - 10^{-5}Y_0^2)$.

In our example $3\omega_{min} > \omega_{max}$, that is, it lies outside the passband of the system, this being the grounds for discarding the third harmonics of the fundamental oscillation in the derivation of the abbreviated equations.

Figure 3 shows the frequency dependence of the amplitude of the oscillations, while Figure 4 shows the frequency dependence of the delay-time curves for this example. It is seen from Figure 4 that in the present example the regions with and without oscillations do not alternate. This means that the second critical value of delay time lies inside the range

of τ corresponding to the first oscillation region, and consequently m = 1. Using (11) we obtain for our example $\tau_{cr2} = 0.85$ microsecond.

3. INVESTIGATION OF STABILITY OF STATIONARY PERIODIC SOLUTIONS

As $\tau \to \infty$, the number of possible oscillations also becomes infinite, occupying the entire frequency spectrum in the passband of the system, while the intervals between frequencies decrease to zero. Not all types of oscillations are stable for a specific value of τ. Many papers [1, 2, 3] deal with the stability of oscillations when several frequencies fall into the passband. It is shown in these papers that the stability conditions are most difficult to satisfy for frequencies lying in the vicinity of the limits of the passband, and easiest to satisfy for frequencies lying near the resonant frequency.

There are no references in the literature to a general investigation of the behavior of oscillating systems with delay time varying over a sufficiently wide range.

The stability-investigation methods used in [1 — 3] do not answer the question of the stability of oscillations near the resonant frequency, and can therefore not be used in our case.

It must be emphasized that the stability-investigation method used in the present section cannot answer in principle the question of the possibility of occurrence of other types of oscillations due to the type of nonlinearity.

Let us put

$$\begin{aligned} Y &= Y_0 + \xi, \\ \psi &= \psi_0 + \varphi = \omega t + \varphi(t). \end{aligned} \tag{13}$$

Here ξ and φ are small deviations from the stationary values and vary slowly in time. The initial conditions can be specified on the interval $-\tau \leqslant t \leqslant 0$ in the form $\xi_\tau = \xi_0$, $\varphi_\tau = \varphi_0$. We also stipulate that $\xi = \varphi = 0$ when $t \leqslant 0$.

We substitute (13) in (3) and use conditions (6); after discarding the higher-order terms we obtain a system of two linear differential equations with delay argument

$$\begin{aligned} -\dot{\xi} T_\kappa &= \xi + K_\xi \cos \omega\tau \xi_\tau - K_Y Y_0 \sin \omega\tau (\varphi - \varphi_\tau), \\ \dot{\varphi} T_\kappa &= \frac{\omega_0^2 - \omega^2}{2\delta\omega Y_0} \xi + \frac{1}{Y_0} K_\xi \sin \omega\tau \xi_\tau + K_Y \cos \omega\tau (\varphi - \varphi_\tau). \end{aligned} \tag{14}$$

Here $T_k = 1/\delta$.

Going to an operator notation and being interested only in stability, we obtain by using the steady-state conditions (5)

$$\begin{aligned} -PT_\kappa \xi(P) &= (1 - \gamma e^{-P\tau}) \xi(P) + Y_0 \frac{\omega_0^2 - \omega^2}{2\delta\omega} (1 - e^{-P\tau}) \varphi(P), \\ PT_\kappa \varphi(P) &= \frac{\omega_0^2 - \omega^2}{2\delta\omega Y_0} (1 - \gamma e^{-P\tau}) \xi(P) - (1 - e^{-P\tau}) \varphi(P). \end{aligned} \tag{15}$$

Here $\gamma = K_\xi / K_Y$.

The stability of the amplitude, as usual, is checked against the sign of the real part of the roots of the determinant

$$D(\lambda) = \begin{vmatrix} \lambda + 1 - \gamma e^{-\lambda\tau/T_\kappa} & Y_0 \frac{\omega_0^2 - \omega^2}{2\delta\omega} (1 - e^{-\lambda\tau/T_\kappa}) \\ \frac{\omega_0^2 - \omega^2}{2\delta\omega Y_0} (1 - \gamma e^{-\lambda\tau/T_\kappa}) & -(\lambda + 1 - e^{-\lambda\tau/T_\kappa}) \end{vmatrix} = 0. \tag{16}$$

We use here the symbol $\lambda = PT_k$.

After making some obvious transformations, (16) assumes the form

$$D(\lambda) = \lambda^2 + 2\lambda + K_Y^2 - (1+\gamma)(\lambda + K_Y^2)\, e^{-\lambda\tau/T_\kappa} + \gamma K_Y^2 e^{-2\lambda\tau/T_\kappa} = 0. \tag{17}$$

Putting $\lambda = j\beta$ we obtain from (17) the equation $D(j\beta) \neq 0$, which determines the mapping of the imaginary axis on the complex D plane. Variation of β causes a variation of the real and imaginary parts of $D(j\beta)$. In the case when $\pm j\beta$ are roots of (17) we have $D(j\beta) = 0$.

The stability of (17) depends on the parameters K_Y, T_k, γ and τ. In the case when the variation of the parameter τ from zero upward results in the appearance of imaginary roots, the system is at the threshold of stability. Upon further variation of τ, it becomes unstable (provided it is stable when $\tau = 0$). Putting $\tau = 0$ in (17) and using the Hurwitz criterion, we can show that the oscillations $\tau = 0$ will be stable (it is implied that the phase relations necessary for oscillation at the investigated frequency remain in force).

Imaginary roots can occur only if the following conditions are satisfied simultaneously

$$\begin{aligned} \operatorname{Re} D(j\beta) &= 0, \\ \operatorname{Im} D(j\beta) &= 0. \end{aligned} \tag{18}$$

Equations (18) can be reduced to the form

$$\frac{\sqrt{\alpha^3 + 3\alpha^2 + 3\alpha + 1}}{\frac{1}{2(1+\gamma)}\alpha^2 + \frac{2+(1+\gamma)^2}{2(1+\gamma)}\alpha + 1} \cos\left[\sqrt{\alpha}\,\frac{\tau}{T_\kappa} + \operatorname{arc\,tg}\sqrt{\alpha}\right] = 1. \tag{19}$$

Here $\alpha = \beta^2/K_Y^2$.

Equation (19) enables us to investigate the appearance of the imaginary roots as the parameter τ is varied. However, if τ is varied continuously, the expression preceding the cosine, which we shall henceforth denote by $W(\alpha)$, will have varying coefficients. To avoid this, we change τ by a discrete amount $\Delta\tau = 2\pi/\omega$. With such a discrete variation of τ, the values of the coefficients in the expression for $W(\alpha)$ will remain constant. If there exists a domain $|W(\alpha)| > 1$, then we can cause Equation (19) to be satisfied by varying τ, and this will correspond to the appearance of a pair of pure imaginary roots.

However, there may not be any domains $|W(\alpha)| > 1$, and then the first-approximation equations will be inadequate for the determination of the stability. For example, $|W(\alpha)| < 1$ for all α in the case when $-1 < \gamma < 0$. Then if $\tau > T_k$ the amplitude of the oscillations will be stable in the small and unstable in the large. Usually in our nonlinearity approximation $\gamma < -1$.

We can determine from (19) a delay time (third critical value τ_{cr3}) such that when the inequality $\tau > \tau_{cr3}$ is satisfied it is impossible to have stationary oscillations of stable amplitude at any frequency, including $\omega = \omega_0$.

To find τ_{cr3} we first determine α_{max} from the condition $|W(\alpha)_{max}| = 1$. In order for Equation (19) to be satisfied in this case, an additional condition must be satisfied, namely: if $W(\alpha)_{max} = 1$, then

$$\tau = \frac{T_\kappa}{\sqrt{\alpha_{max}}}(2\pi n - \operatorname{arc\,tg}\sqrt{\alpha_{max}}), \qquad n = 1, 2, \ldots. \tag{20}$$

On the other hand, if $W(\alpha_{max}) = -1$, then

$$\tau = \frac{T_\kappa}{\sqrt{\alpha_{max}}}[(2n-1)\pi - \operatorname{arc\,tg}\sqrt{\alpha_{max}}], \qquad n = 1, 2, \ldots. \tag{21}$$

The value of τ_{cr3} is accordingly determined from (20) or (21) in which we put $n = 1$. In our calculations we take the positive root in the expression for $W(\alpha)$.

Figure 5 shows plots which illustrate the foregoing statements.

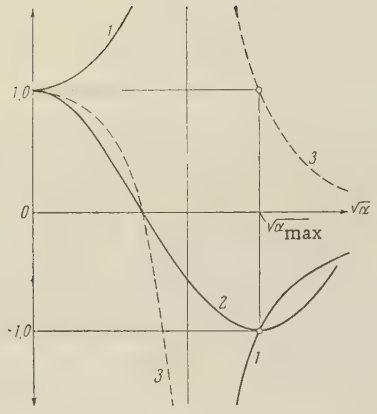


Figure 5. Curves for the determination of the stability of stationary oscillations:

1 — $W(a)$; 2 — $\cos\left(\frac{\sqrt{\alpha\tau}}{T_\kappa} + \text{arctg}\sqrt{\alpha}\right)$;

3 — $W(a)\cos\left(\frac{\sqrt{\alpha\tau}}{T_\kappa} + \text{arctg}\sqrt{\alpha}\right)$

If $\omega = \omega_0$ we get from (16)

$$(\lambda + 1 - K_\xi e^{-\lambda\tau/T_\kappa})\,[\lambda + 1 - e^{-\lambda\tau/T_\kappa}] = 0. \qquad (22)$$

The expression in the square brackets cannot vanish for any value of λ, with the exception of $\lambda = 0$. Consequently

$$D(\lambda) = \lambda + 1 - K_\xi e^{-\lambda\tau/T_\kappa} = 0. \qquad (23)$$

Expression (23) can have roots with positive real part only when $|K_\xi| > 1$.

Substituting $\lambda = j\beta$ in (23) and equating the re and imaginary parts of the resultant expression t zero, we get

$$1 - K_\xi \cos \beta\tau/T_\kappa = 0,$$
$$\beta + K_\xi \sin \beta\tau/T_\kappa = 0. \qquad (24)$$

The same expressions can be written in a different fashion

$$W(Y_0, \xi, \beta) = \frac{|K_\xi|}{\sqrt{\beta^2 + 1}} = 1, \qquad (25a)$$

$$\tau = \frac{T_\kappa}{\beta}\,[2\pi n - \text{arc tg}\,\beta] \quad \text{for} \quad K_\xi > 1, \quad n = 1, 2, \ldots, \qquad (25b)$$

$$\tau = \frac{T_\kappa}{\beta}\,[(2n - 1)\,\pi - \text{arc tg}\,\beta] \quad \text{for} \quad K_\xi < -1, \; n = 1, 2, \ldots. \qquad (25c)$$

In our notation, β is a dimensionless frequency. To go over to a dimensional frequency, we must use the expression $\Omega = \beta/T_\kappa$; $W(Y_0, \xi, \beta)$ is a transfer function of the envelope perturbation, for the closed feedback loop of the oscillator.

Example

$$K_Y = R\,(10^{-3} - 10^{-5}\,Y_0^2),$$
$$K_\xi = R\,(10^{-3} - 3\cdot 10^{-5}\,Y_0^2),$$
$$\omega = \omega_0, \qquad R = 2.5\,\text{Kohm}$$
$$T_\kappa = 50\ \mu\text{sec}$$

Using the steady-state conditions (6), we find that $K_\xi = -2$. Then the dimensionless frequency is $\beta = 1.73$, hence $\Omega = 3.46 \cdot 10^3$ sec. From (25c) with n = 1 we obtain $\tau_{cr3} = 61$ microseconds.

We have shown thus that no stable oscillations can be produced in principle in an oscillator when $\tau > \tau_{cr3}$, for even in the case when $-1 < \gamma < 0$ there can always be found, for large delay, a sufficiently strong perturbation which leads to the region of values $|\gamma| > 1$.

When $\tau \simeq \tau_{cr3}$ the delay becomes so large that a noticeable phase shift is produced in the envelope of the high-frequency oscillations; the additional phase shift in the envelope is introduced by the resonant circuit; in addition, the form of nonlinearity can also influence the phase of the envelope. In the case when the total phase advance of the envelope is 2π, conditions are created for oscillations.

When τ is slightly greater than τ_{cr}, sinusoidal self-modulation of insignificant depth sets in. The period of the modulation is not a multiple of the delay time, owing to the considerable influence exerted by the resonant circuit on the phase of the envelope. Self-modulation is accompanied by phase modulation of the carrier.

Further increase in τ increases the depth of modulation and distorts the form of the envelope. The period of the modulation becomes a multiple of τ. It can be expected that at even greater delay times the amplitude of the high-frequency oscillations can drop to zero at individual instants of time. Consequently, individual pulses of high-frequency oscillations can occur. Since small random perturbations are always acting in the system, the position of the pulse front will vary all the time, and the pulse duration will be a random quantity with their tops having random crests and trough. Any periodicity in the alternation of the individual bursts will disappear. Oscillations which are separated by an interval will be found to be uncorrelated not only relative to the high-frequency component, but also relative to the envelope. An oscillator with delay thus becomes, at large dleay times, a noise generator, generating the entire frequency spectrum in the passband of the system.

Reference [4] describes an experiment which confirms to some degree the last assumption.

REFERENCES

1. I.S. Gonorovskiy, On the theory of high-frequency oscillators with time-delay feedback, Radiotekhnika, 1958, 13, 5, 19.
2. V. Met, On multimode oscillators with constant time delay, Proc. I.R.E., 1957, 45, 8, 1119.
3. W.A. Edson, Frequency memory in multimode oscillators, I.R.E. Trans. Circuit Theory, 1955, CT-2 1, 58.
4. O.E. Lange, Experiments on the regeneration of binary microwave pulses, Bell System Techn. J. 1956, 35, 1, 67.

Received by the editors 1 June, 1960
after revision 7 February, 1961

SYNCHRONIZATION OF AN OSCILLATOR BY A WEAKLY MODULATED EXTERNAL SIGNAL

V. P. Yakovlev

We consider the synchronization of an oscillator by an amplitude and frequency modulated external voltage under the assumption that the phase difference between the oscillations of the oscillator and of the external voltage, and also the amplitude of the oscillator, are close to the corresponding values obtained when the EMF is not modulated. The dependence of the depth of modulation of the oscillator on the frequency of the voltage modulation and on the deviation of the frequency of the oscillations from the fundamental frequency was investigated for amplitude-modulated, frequency-modulated, and biharmonic external signals. The results obtained are in qualitative agreement with the theoretical calculation.

1. FORMULATION OF PROBLEM AND INITIAL EQUATIONS

The action of an external EMF containing more than one frequency component on an oscillator in the self-excitation mode has been treated in several articles. In [1 — 4] the synchronization of an oscillator by a sequence of radio pulses was investigated, and the establishment

of the amplitude and phase of its oscillations were considered. References [5, 6] take account of the presence of fluctuation noise. The influence of the modulation of the EMF on the synchronization of the oscillator has been considered in [7—11]. In particular, [11] reports an investigation of synchronization by means of an external EMF which is modulated in amplitude and in frequency. The synchronization band is calculated and an expression derived for the phase of the oscillator under the assumption that the phase difference between the oscillations and the external EMF is small.

In the present paper we investigate the behavior an an oscillator under the influence of an external EMF which is modulated in amplitude and in frequency, with a sufficiently small depth of modulation. This condition does not mean that the phase difference between the oscillations and the external EMF is small, but rather it is close to the steady-state value produced by the unmodulated EMF. The amplitude of the oscillations also differs little from the steady-state value.

Assume that the external EMF applied to the grid of the tube of the oscillator is

$$E(t) = E_0 (1 + m_1 \cos \Omega_1 t) \cos [\omega t + \varphi (\Omega_2 t)], \tag{1}$$

where

$$\frac{d\varphi}{dt} = m_2' \omega \sin \Omega_2 t; \tag{2}$$

m_1 and m'_2 are the coefficients of amplitude and frequency modulation, respectively.

We approximate the dependence of the plate current on the grid voltage v by a third-degree polynomial

$$I_a (v) = Sv - \gamma' v^3. \tag{3}$$

Then the equation for the dimensionless voltage x on the grid can be written

$$\frac{d^2x}{dt'^2} + x = 2\mu \left[x + \xi x - 4x^2 \frac{dx}{dt'} + A_0 (1 + m_1 \cos \gamma_1 \tau) \cos \{t' + \varphi (\gamma_2 \tau)\} \right]. \tag{4}$$

where A_0 is the dimensionless amplitude of the external EMF, μ a small parameter (increment of the oscillator), ξ — dimensionless detuning, defined by the relation

$$2\mu\xi = 1 - \frac{\omega_0^2}{\omega^2} \simeq \frac{2(\omega - \omega_0)}{\omega}; \tag{5}$$

ω_0 — frequency of oscillation, $t' = \omega t$ — dimensionless time; $\tau = \mu t'$ — so-called "slow-time"; the quantities γ_1 and γ_2 are defined by the expressions

$$\mu\gamma_1 = \frac{\Omega_1}{\omega}; \qquad \mu\gamma_2 = \frac{\Omega_2}{\omega}. \tag{6}$$

Taking account of the fact that the Q of the circuit is high ($\mu \ll 1$), we can seek a solution by the method of slowly varying amplitude ρ and phase θ, putting $x = \rho(\tau) \cos [t' + \Theta(\tau)]$. The usual procedure then results in the following van der Pol equations for the amplitude ρ and the difference α between the phase of the oscillator θ and the phase of the signal EMF φ:

$$\begin{aligned} \rho' &= \rho (1 - \rho^2) - A_0 (1 + m_1 \cos \gamma_1 \tau) \sin \alpha, \\ \rho\alpha' &= -\xi\rho - A_0 (1 + m_1 \cos \gamma_1 \tau) \cos \alpha - \rho m_2 \sin \gamma_2 \tau, \end{aligned} \tag{7}$$

where

$$m_2 = \frac{m_2'}{\mu}. \tag{8}$$

An exact solution of (7) cannot be obtained in the general case, nor is a qualitative investigation possible owing to the explicit dependence on τ. We therefore confine ourselves to a solution of the linearized equations, taking as the zero-order approximation the steady-state solution in the absence of modulation ($m_1 = m'_2 = 0$), that is, the solution of the equations

$$\rho'_0 = \rho_0(1 - \rho_0^2) - A_0 \sin\alpha_0 = 0,$$
$$\rho_0\alpha'_0 = -\xi\rho_0 - A_0\cos\alpha_0 = 0. \tag{9}$$

It is natural to expect that at small m_1 and m_2 the solution of (7) will be close to the solution of (9). Moreover, even if m_1 and m_2 cannot be regarded as small, but the frequencies γ_1 and γ_2 are sufficiently large, the amplitude and the phase of the oscillations will not be able to follow the variations in the amplitude and phase of the external signal, and consequently, the solution of (7) should in this case differ little from ρ_0 and α_0. We can thus count on the solution of (7) in the form

$$\rho = \rho_0 + \rho_1, \quad \alpha = \alpha_0 + \alpha_1, \tag{10}$$

where ρ_1 and α_1 are small enough ($\rho_1 \ll \rho_0$, $\alpha_1 \ll \pi/2$), to be suitable for the two foregoing cases.

2. SOLUTION OF THE FIRST-APPROXIMATION EQUATIONS

We start from the assumption that m_1 and m_2 are small. Substituting (10) in (7) and confining ourselves to terms linear in ρ_1, α_1, m_1, and m_2, we obtain the following equations for ρ_1 and α_1:

$$\rho'_1 + (3\rho_0^2 - 1)\rho_1 - \xi\rho_0\alpha_1 = -\rho_0(1 - \rho_0^2)m_1\cos\gamma_1\tau,$$
$$\alpha'_1 + \frac{\xi}{\rho_0}\rho_1 + (\rho_0^2 - 1)\alpha_1 = \xi m_1\cos\gamma_1\tau - m_2\sin\gamma_2\tau. \tag{11}$$

Solving this equation, we obtain expressions for the amplitude ρ and for the frequency $\omega(t)$ of the oscillator

$$\rho = \rho_0[1 + M_{11}\cos(\Omega_1 t + \psi_1) - M_{12}\cos(\Omega_2 t + \psi_2)],$$
$$\omega(t) = \frac{d}{dt}[\omega t + \Theta(t)] = \omega[1 + M_{21}\sin(\Omega_1 t + r_1) + M_{22}\sin(\Omega_2 t + r_2)], \tag{12}$$
$$M_{11} = m_1R_1; \quad M_{12} = m_2R_2; \quad M_{21} = \mu m_1Q_1; \quad M_{22} = m'_2Q_2.$$

The expressions for R_1, R_2, Q_1, and Q_2 are rather cumbersome and will not be given here. The values of R_1, R_2, Q_1 and Q_2 depend on the amplitude ρ_0, on the detuning ξ, and on the modulation frequencies γ_1 and γ_2. To illustrate this dependence, we choose $A_0^2 = 0.1$, corresponding to a small external signal ($A_0^2 < 8/27$) [12].

Figure 1 shows the dependence of ρ_0^2 on ξ (solid curve). The half-width of the synchronization band at $A_0^2 = 0.1$ corresponds to $\xi = 1.32$.*

The solid lines of Figures 2 — 5 are plots of $R_1(\gamma_1)$, $R_2(\gamma_2)$, $Q_1(\gamma_1)$ and $Q_2(\gamma_2)$ at certain values of the detuning ξ, as functions of $\log_{10}\gamma^2$, and, as before, we put

$$A_0^2 = 0.1.$$

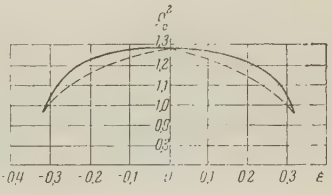


Figure 1. Dependence of the oscillation amplitude on the detuning.

At small detunings ξ, we can obtain from (11) simple approximate formulas for the values of Q_1 and Q_2 involved in (12). If we assume, as is done in (11), that the detuning ξ and the change in amplitude ρ_1 are so small that we can neglect $\xi\rho_1$ in the second equation

*By "synchronization interval" we mean here and throughout the interval obtained in response to an unmodulated external EMF.

of (11), we obtain the following equation for the phase

$$\alpha_1' + (\rho_0^2 - 1)\,\alpha_1 = \xi m_1 \cos \gamma_1 \tau - m_2 \sin \gamma_2 \tau, \tag{13}$$

solution of which yields

$$Q_1 = \frac{\xi \gamma_1}{\sqrt{(\rho_0^2 - 1)^2 + \gamma_1^2}}, \qquad Q_2 = \frac{\rho_0^2 - 1}{\sqrt{(\rho_0^2 - 1)^2 + \gamma_2^2}}. \tag{14}$$

A comparison of (12) and (14) shows that the error in the calculation of Q_1 and Q_2 by means of formulas (14) does not exceed 10 — 15 percent if $\xi < 0.22$.

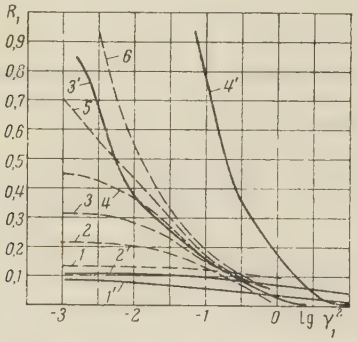


Figure 2. Amplitude modulation of the oscillator in the case of amplitude modulation of the external signal

1 — ξ = 0; 2 — ξ = 0.2; 3 — ξ = 0.28; 4 — ξ = 0.29; 5 — ξ = 0.3; 6 — ξ = 0.315; 1′ — ξ = 0; 2′ — ξ = 0.22; 3′ — ξ = 0.31; 4′ — ξ = 0.32

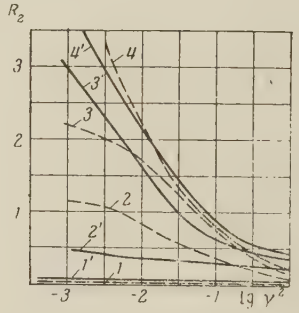


Figure 3. Amplitude modulation of the oscillator in the case of frequency modulation of the external signal.

1 — ξ = 0; 2 — ξ = 0.2; 3 — ξ = 0.29; 4 — ξ = 0.315; 1′ — ξ = 0.1; 2′ — ξ = 0.22; 3′ — ξ = 0.31; 4′ — ξ = 0.32

At even smaller detunings ξ, we find from (9) that $\cos \alpha_0 \simeq 0$ and $\rho_0^2 - 1 \simeq A_0/\rho_0$. Substituting this value of $\rho_0^2 - 1$ into (14) we obtain for Q_1 and Q_2 expressions which are analogous to Equations (14) and (21) of [11]:

$$Q_1 = \frac{\xi \gamma_1}{\sqrt{\gamma_1^2 + \frac{A_0^2}{\rho_0^2}}}, \qquad Q_2 = \frac{A_0}{\rho_0} \frac{1}{\sqrt{\gamma_2^2 + \frac{A_0^2}{\rho_0^2}}}. \tag{15}$$

It follows from a comparison of (15) with (14) that the formulas in (15) can be used only when $\xi < 0.1$.

According to (12) the quantities R_1 and Q_1 characterize the modulation of the oscillations upon synchronization by an amplitude-modulated external signal ($m_2 = 0$). The quantities R_2 and Q_2 describe the modulation of the oscillations in the case of frequency modulation ($m_1 = 0$). In the case when the synchronizing voltage is biharmonic, that is

$$E(t) = E_0 [\cos \omega t + m \cos (\omega + \Omega) t], \tag{16}$$

with $m \ll 1$, the expressions for ρ and $\omega(t)$ can also be written in the form

$$\begin{aligned} \rho &= \rho_0 [1 + M_1 \cos (\Omega t + \beta_1)], \\ \omega(t) &= \omega [1 + M_2 \cos (\Omega t + \beta_2)], \\ M_1 &= P_1 m_1, \qquad M_2 = \mu m P_2, \end{aligned} \tag{17}$$

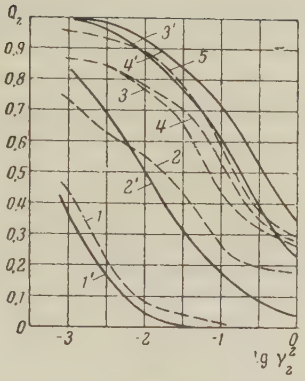


Figure 4

and the values of P_1 and P_2 are obtained from R_1, R_2, Q_1, and Q_2 with $\gamma_1 = \gamma_2 = \gamma$, while γ is given by the expression $\mu\gamma = \Omega / \omega$.

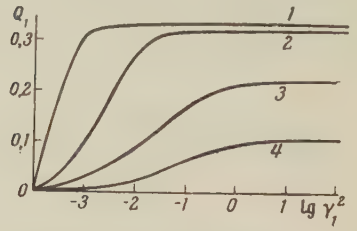


Figure 5

Figure 4. Frequency modulation of the oscillator in the case of frequency modulation of the external signal. $1 - \xi = 0.315$; $2 - \xi = 0.29$; $3 - \xi = 0.2$; $4 - \xi = 0$; $5 - \xi = 0.15$; $1' - \xi = 0.319$; $2' - \xi = 0.31$; $3' - \xi = 0.22$; $4' - \xi = 0$; $5' - \xi = 0.1$

Figure 5. Frequency modulation of the oscillator in the case of amplitude modulation of the external signal. $1 - \xi = 0.319$; $2 - \xi = 0.31$; $3 - \xi = 0.218$; $4 - \xi = 0.1$

Figures 6 and 7 show (solid curves) plots of $P_1(\gamma)$ and $P_2(\gamma)$ for several values of ξ and for $A_0^2 = 0.1$.

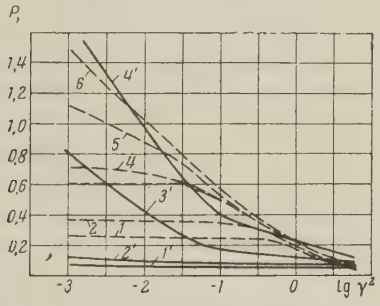


Figure 6. Amplitude modulation of the oscillator in the case of biharmonic signal. $1 - \xi = 0$; $2 - \xi = 0.15$; $3 - \xi = 0.25$; $4 - \xi = 0.28$; $5 - \xi = 0.3$; $6 - \xi = 0.315$; $1' - \xi = 0$; $2' - \xi = 0.22$; $3' - \xi = 0.31$; $4' - \xi = 0.32$

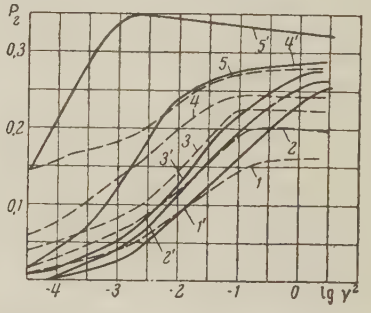


Figure 7. Frequency modulation of the oscillator in the case of biharmonic signal. $1 - \xi = 0$; $2 - \xi = 0.2$; $3 - \xi = 0.28$; $4 - \xi = 0.3$; $5 - \xi = 0.315$; $1' - \xi = 0$; $2' - \xi = 0.1$; $3' - \xi = 0.22$; $4' - \xi = 0.31$; $5' - \xi = 0.319$

3. EXPERIMENT

To ascertain the degree of suitability of the idealization assumed in the theoretical calculation, we investigated experimentally synchronization of an oscillator by means of amplitude-modulated, frequency-modulated, and biharmonic voltages, with low modulation. In the investigation we used a setup whose principal elements are shown in Figure 8. The fundamental frequency ω of the external signal was chosen such as to make the difference between ω and the oscillator frequency ω_0 lie inside the synchronization interval. The value of

$\omega - \omega_0$ was changed by retuning the tank circuit of the oscillator. The detuning $\omega - \omega_0$ was maintained constant during each measurement. The depths of the frequency and amplitude modulations of the oscillator were measured with the aid of amplitude and frequency detectors, an output low-frequency amplifier, and a voltmeter.

In all types of external-signal modulation we investigated the amplitude and frequency modulation of the oscillations as functions of the external-signal modulation frequency at several values of the detuning $\omega - \omega_0$. The amplitude of the synchronizing EMF was chosen such as to make $A_0^2 = 0.01$.

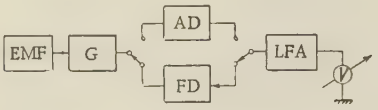


Figure 8. Block diagram of experimental setup.

EMF — source of external synchronizing signal: G — investigated oscillator; AD — amplitude detector; FD — frequency detector; LFA — low-frequency output amplifier; V — voltmeter.

The source of amplitude-modulated signals was a standard signal generator GSS - 6, while the biharmonic EMF was produced by two such generators. The frequency-modulated signals were generated by a Colpitts oscillator using two semiconductor diodes for the variable capacitance in the tank circuit.

It was natural to investigate first the synchronization of the oscillator by means of an unmodulated voltage.

The measured detuning $\omega - \omega_0$ corresponding to the boundary of the synchronous mode was 6 kc. Substituting this value of $\omega - \omega_0$ into (5) and setting $\xi = 0.32$, we can find that μ is equal to 0.02. Figure 1 shows together with the theoretical curve of $\rho_0(\xi)$ the experimental plot (dashed) of the oscillator amplitude ρ_0 as a function of the detuning ξ.

To be able to compare directly the experimental data with the theoretical result it is advantageous to choose an oscillator circuit in which the differential equation for the grid voltage coincides with Equation (4). This requirement is met by the circuit shown in Figure 9. The plate-tuned oscillator uses a 6Zh3T tube, with stabilized + 105 volts DC applied to the screen grid and to the plates. The feedback voltage was applied to grid g_1 from the plate of a 6N1P tube. This grid was maintained at zero potential relative to the cathode by means of resistances R_1 and R_2 and the voltmeter V. The natural frequency of the oscillator was in the 900 kc band. The measured Q of the tank circuit was 100.

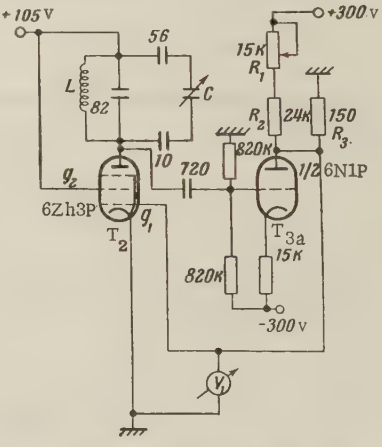


Figure 9. Principal circuit of the oscillator.

The source of modulating voltage for amplitude and frequency modulation was a spectrum and frequency characteristic analyzer ASChKh-1; the output of this instrument, a voltage whose frequency was automatically varied from 0 to 20 kc, was measured with a frequency meter. This made it possible to obtain, at the output of the sources of the two types of modulation, weakly modulated signals with variable modulation frequency; these signals synchronized the investigated oscillator. The detected and amplified low-frequency signals were measured with voltmeter V and fed to the input of the ASChKh-1. Thus, it was possible to observe on the screen of this instrument the dependence of the coefficients of modulation of the oscillator on the modulation frequency. The frequency meter and voltmeter made it possible to establish this dependence quantitatively.

The use of the ASChKh-1 instrument eliminated errors in the measurement of the depth of modulation of the oscillator; such errors are due to slow changes in the detuning in the measurement, owing to the instability of the voltage sources.

To eliminate these errors, use was made of the slowness of the variation of the detuning; the modulation frequency changed from zero to maximum so rapidly, that the detuning remained almost constant during this time.

The coefficient m'_2 of the depth of frequency modulation was chosen equal to $3 \cdot 10^{-4}$, thus ensuring sufficiently small values of m_2. The depth of amplitude modulation was 0.02.

Figures 2 — 4 show, along with the theoretical curves, also the results of the measurement of R_1, R_2, and Q_2 (dashed curves).

In measuring the frequency modulation of an oscillator synchronized by amplitude-modulated signals, it was found that this frequency modulation is comparable with the parasitic frequency modulation inherent in the source. This prevented a direct measurement of the coefficient M_{21} when working with this source. Nevertheless, it proved possible to estimate the upper limit of M_{21} and to establish that $M_{21} < \mu m_1$,

In the case of a biharmonic voltage, the value of m as given by (17) was chosen to be 0.04 The difference frequency was measured with a frequency meter connected to the output of the amplitude-modulation detector. Figures 6 and 7 show the results of the measurement of P_1 and P_2 (dashed curve) .

4. ANALYSIS OF RESULTS AND CONCLUSIONS

A comparison of the experimental results with the theoretical ones, which can be done by using the plots of Figures 1 — 6, shows that they are in qualitative agreement. The qualitative discrepancies, which can be seen on all the plots, are apparently due to the inaccuracy in the cubic approximation of the plate current vs. grid voltage dependence. These discrepancies are particularly strong in the case of amplitude modulation at small detunings, that is, $\xi \simeq 0$. As can be seen from the plots of Figures 2, 3, and 6, the experimentally measured depth of modulation is two or three times greater in this case than the theoretically calculated value.

When the oscillator is synchronized by a frequency modulated voltage, the depth of frequency modulation of the oscillations M_{22}, is close to the depth of modulation of the signal at small values of the detuning ξ and at low modulation frequencies (Figure 4) . As ξ is increased the frequency modulation decreases. With increasing modulation frequency, the coefficient M_{22} decreases monotonically, Amplitude modulation is low when $\xi \simeq 0$, and increases strongly for small modulation frequencies, when $\xi - x_i$ approaches the limit of the synchronization interval (Figure 3) . If the modulation frequency increases, the amplitude modulation decreases and practically disappears at $\gamma \geqslant 1$.

If an amplitude-modulatied EMF is applied to the oscillator, then, as can be seen from Figure 2, the amplitude modulation of the oscillator is small for small detunings ξ. As ξ approaches the limit of the synchronization interval, if the modulation frequencies are low, the depth of amplitude modulation increases strongly. With increasing modulation frequency, the amplitude modulation decreases and is practically insignificant when $\gamma \geqslant 1$. The frequency modulation is quite small for all values of ξ and Ω .

Synchronization by means of a biharmonic EMF (Figures 6 and 7) is accompanied by amplitude modulation, which is relatively small when $\xi \simeq 0$ and increases strongly at low modulation frequencies and as ξ approaches the limit of the synchronization interval. The frequency modulation is insignificant, and unlike the amplitude modulation it increases with increasing modulation frequency.

Let us compare in conclusion the influence of amplitude and frequency modulations of the signal voltage on the modulation of the oscillator. Figures 4 and 4 show that Q_1 does not exceed unity, while Q_2 is close to unity at small ξ. Therefore, as follows from (12), when $m_1 \simeq m'_2$ the frequency modulation of the signal brings about a considerably greater modulation of the oscillator frequency, than an amplitude modulation of the signal with the same depth. An analogous conclusion can be made by comparing the influence of both types of signal modulation on the amplitude modulation of the oscillator: at comparable depths of signal modulation ($m_1 \simeq m'_2$) the frequency modulation of the signal has a much stronger effect on the oscillator.

I consider it my pleasant duty to express deep gratitude to S.M. Rytov and F.A. Vodop'yanov for help with the work and for valuable remarks.

REFERENCES

1. P.N. Zanadvorov, Radiotekhnika i Elektronika, 1958, 3, 2, 202.
2. E.S. Voronin, R.G. Khokhlov, Radiotekhnika i Elektronika, 1956, 1, 1, 79.
3. E.S. Voronin, G.N. Berestovskiy, Radiotekhnika, 1956, 11, 3, 34.
4. E.S. Voronin, I.I. Rogatnev, Radiotekhnika i Elektronika, 1957, 2, 2, 144.

5. S. Ya. Rayevskiy, R.G. Khokhlov, Radiotekhnika i Electronika, 1958, 3, 4, 507.
6. Yu.F. Korobov, Trudy. LEIS, (Transactions, Leningrad Electrotechnic Institute for Communications) 1959, No. VII (44), 59.
7. D.P. Martynenko, R.G. Khokhlov, Radiotekhnika i Elektronika, 1957, 2, 8, 1001.
8. I.M. Sobol, Radiotekhnika i Elektronika, 1959, 4, 10, 1739.
9. A.N. Shusterovich, Radiotekhnika i Elektronika, 1960, 5, 4, 568.
10. Yu.I. Samoilenko, Radiotekhnika, 1960, 15, 7, 37.
11. E.S. Voronin, Radiotekhnika, 1959, 14, 2, 48.
12. A.A. Andronov, Collected Works, Izd. AN SSSR, 1956.

Submitted to the editor, 2 January, 1961

OSCILLATOR PARAMETRICALLY COUPLED WITH A LINEAR NETWORK

Yu. V. Grigor'yev and R.V. Khokhlov

The article is devoted to a study of parametric excitation of oscillations in a linear system, taking into account the reaction of the excited oscillations on the apparatus that effects the chang in the energy-storage parameter. The analysis of the problem reduces to an investigation of an oscillating system with two parametrically coupled degrees of freedom. Stationary oscillations in such a system are studied and their stability analyzed.

INTRODUCTION

It is known that in parametric resonance the presence of damping does not limit the increas in the amplitude of the oscillations excited in a linear system [1]. If the system can be regarde as linear up to sufficiently large amplitudes, the factor limiting the increase in the amplitude is the reaction of the system on the generator which produces the change in the energy storage parameter. Obviously such a reaction becomes noticeable when the oscillating power in the system is of the same order of magnitude as the generator power.

In the case of ordinary resonance, the problem of taking into account the reaction on the generator reduces to an investigation of an oscillator coupled to the network as a single oscillating system with two degrees of freedom. This investigation has shown [2, 3, 4] that when the coupling is much less than a certain critical value, the power dissipated in the resonant circuit is much less than the generator power, and the limit on the oscillation amplitude in the resonant circuit is imposed by the attenuation in it. On the other hand, if the coupling becomes of the same order as the critical value, the oscillation power in the resonant circuit is of the same order as the generator power, and the reaction brings about a noticeable change in amplitude and frequency of the generator. These changes are the factors that limit the increase in the amplitude in the resonant circuit.

To solve the corresponding problem for parametric resonance, it is necessary to investigate an oscillator which is parametrically coupled with a linear network.

The simplest example of parametric coupling between systems is the coupling between vertical oscillations of a spring pendulum and its angular oscillations, when the frequency of the vertical oscillations is twice the frequency of the angular oscillations [5]. If linear systems are coupled parametrically, then, as in the case of direct coupling between systems, the action

of oscillations in one partial system on the oscillations in the second partial system leads to periodic pumping of energy from one system to the other. The increase in the amplitude in each of the systems is limited by the total energy stored in the systems.

A.N. Charakhch'yan [6] considered parametric excitation of oscillations, in which the external voltage with specified amplitude and frequency was fed to a tuned amplifier parametrically coupled with a linear network. In this case the reaction of the excited oscillations on the amplifier limits the growth of the amplitude in the resonant circuit. The frequency of oscillations remains unchanged in view of the fact that the amplifier serves to decouple the network from the generator. The reaction of the oscillations in parametric amplifiers on the generator which changes the parameter is analyzed in [7].

The present paper is devoted to an analysis of the reaction of a parametrically excited network on the generator, when the variation of the parameter is produced by the generator directly, and not through an amplifier as in [6[. The analysis is based on a consideration of an oscillating system with two parametrically coupled degrees of freedom.

1. ABBREVIATED EQUATIONS. STATIONARY SOLUTIONS

The problem posed can be investigated for the system shown in Figure 1, if the oscillator modulated the inductance by means of magnetic nonlinear elements such as ferrites, or for the system of Figure 2, if the oscillator modulates a capacitance by using the nonlinear dependence of the charge on the junction of a semiconductor diode on the applied voltage.

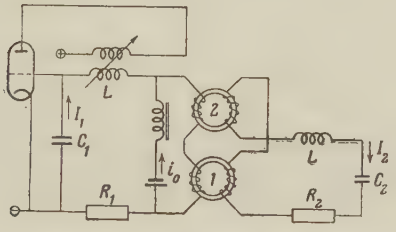


Figure 1. System with parametric coupling produced with ferrite toroids.

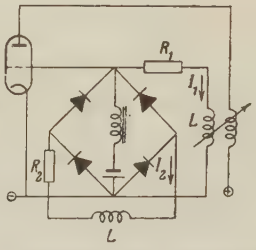


Figure 2. System with parametric coupling produced by semiconductor diodes.

Let us examine the system of Figure 1.

Assuming the toroids to be identical, the toroidal windings 2 to be connected in opposition, and the number of turns to be the same in all windings, we can represent the dependence of the magnetic fluxes in the toroids on the current flowing in the windings in the form

$$\Phi_1 = \Phi(I_1 + I_2), \qquad \Phi_2 = \Phi(I_1 - I_2).$$

By approximating Φ with the series

$$\Phi(I) = \Phi_0 + \beta I + \gamma I^2,$$

we can obtain equations describing the behavior of the system

$$\begin{aligned} &\ddot{X}_1 + 2m\ddot{X}_1\dot{X}_1 + 2\xi^4 m\ddot{X}_2\dot{X}_2 + 2\delta(X_1)\dot{X}_1 + \nu_1^2 X_1 = 0, \\ &(1 + 2m\dot{X}_1)\ddot{X}_2 + 2m\ddot{X}_1\dot{X}_2 + 2\delta_2\dot{X}_2 + \nu_2^2 X_2 = 0. \end{aligned} \tag{1}$$

Here $m = 2\gamma / L_0$. C_1 is the coefficient of parametric coupling; $\xi = \nu_1/\nu_2$; $\nu_1^2 = 1/L_0C_1$; $L_0 = L + 2B$; X_i — voltage on the corresponding capacitance. The remaining notation is standard.

The second equation of (1) differs from the equations for a network in which the parameter is modulated by a source of unlimited power in that the external signal which modulates the parameter in this network depends on the oscillations excited in the network. Excitation of parametric oscillations in a network, as is well known, takes place when $\xi = \nu_1/\nu_2 \simeq \simeq 2; 1; 2/3; 1/2; \ldots$ and at sufficient depth of modulation.

We note that the fact that we confine ourselves to second-order terms in the expansion for the magnetic flux enables us to trace the reaction of the network on the generator in the case when $\xi \simeq 2$. This case will indeed be considered henceforth. In order to trace the reaction on the generator for other values of ξ, it is necessary to include the corresponding terms in the expansion for the magnetic flux.

The harmonic approximation of the solution of system (1) can be sought in the form

$$X_1 = A\cos(\nu_1 t + \varphi), \quad X_2 = B\cos\left(\frac{\nu_1}{2}t + \psi\right),$$

where A, B, φ, and ψ are slowly varying functions of the time.

By the usual methods w e can obtain the following abbreviated equations

$$\begin{aligned} \dot{A} &= \delta(A)A + m\nu_1^2 B^2 \cos\Phi, \\ \dot{B} &= -\delta_2 B - \frac{1}{4}m\nu_1^2 AB\cos\Phi, \\ \dot{\varphi} &= m\nu_1^2 \frac{B^2}{A}\sin\Phi, \\ \dot{\psi} &= \Delta + \frac{1}{4}m\nu_1^2 A\sin\Phi. \end{aligned} \tag{2}$$

Here $\Phi = 2\psi - f$; $\Delta = \nu^2 - \nu_1/2$; $\delta(A)$ is the average increment. From the last two equations of (2) we obtain the following equation for the behavior of the phase difference Φ.

$$\dot{\Phi} = 2\Delta + \left(\frac{A}{2} - \frac{B^2}{A}\right)m\nu_1^2\sin\Phi. \tag{3}$$

The abbreviated equations describing the behavior of the system shown in Figure 2 are analogous in form.

Proceeding to an investigation of the steady-state oscillations in the system, we note that in the stationary states the phase difference Φ is constant, whereas the phases φ and ψ themselves, varying linearly in time, characterize the phase shift due to the reaction of the loading network on the generator:

$$\dot{\varphi} = 2\dot{\psi} = \omega - \nu_1.$$

Taking this fact into account, we obtain for a soft oscillation mode, when $\delta(A) = \delta_0\left(1 - \frac{A^2}{A_0^2}\right)$, where A_0 is the amplitude of the unloaded generator, the following equation for the oscillation frequency:

$$\left(\nu_2 - \frac{\omega}{2}\right)^2\left[1 - \frac{2\delta_2}{\delta_0} - \frac{\delta_2^2}{q^2} - \frac{\left(\nu_2 - \frac{\omega}{2}\right)^2}{q^2}\right]^2 - \left(\frac{2\delta_2}{\delta_0}\right)^2\Delta^2 = 0,$$

where

$$q = \left(\frac{A_0}{4}m\nu_1^2\right),$$

and the equation for the "resonant" curves

$$\eta^2 = \left(1 - \frac{\delta_2^2}{q^2} - y\right)\left(\frac{2\delta_2}{\delta_0} - y\right)^2, \tag{4}$$

where

$$\eta^2 = \left(\frac{2\delta_2}{\delta_0}\right)^2\frac{\Delta^2}{q^2}; \qquad y = 2\frac{2\delta_2}{\delta_0}\frac{B^2}{A^2} = 1 - \frac{A^2}{A_0^2}.$$

The family of amplitude and frequency curves for the case $2\delta_2/\delta_0 = 1/4$ is shown in Figures 3 and 4.

In the case $\frac{2\delta_2}{\delta_0} \geqslant 1 - \frac{\delta_2^2}{q^2}$, the amplitude curves cross the axis $\Delta = 0$ at the point $y = 1 -$

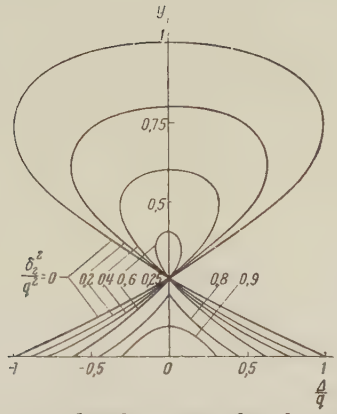


Figure 3. Amplitude curves for the case $2\delta_2/\delta_0 = 1/4$.

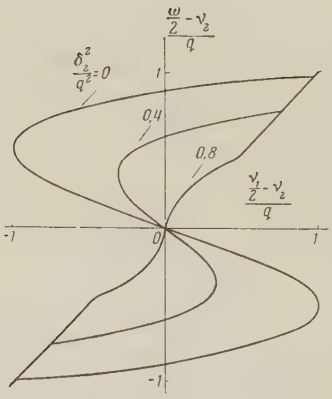


Figure 4. Frequency curves for the case $2\delta_2/\delta_0 = 1/4$.

The amplitude and frequency curves are single-valued here.

In the case when $\frac{2\delta_2}{\delta_0} < 1 - \frac{\delta_2^2}{q^2}$, the axis $\Delta = 0$ is crossed at the points $y = 2\delta_2/\delta_0$ and $y = 1 - \delta_2^2/q^2$. In this case the amplitude and frequency curves become multiple-valued, and this leads to a qualitative change in the behavior of the system, as characterized by the pulling phenomenon.

2. STABILITY OF STATIONARY SOLUTIONS. CASE OF WEAK, MEDIUM, AND STRONG LOADING OF THE GENERATOR

The stability conditions obtained by the Lyapunov method, have the following form for the sought oscillations mode:

$$y < \frac{1}{2}\left(1 + \frac{\delta_2}{\delta_0}\right),$$

$$y^4 - \left(1{,}5 - 2\frac{\delta_2}{\delta_0} + 2\frac{\delta_2^2}{q^2}\right)y^3 + \left(0{,}5 - 3\frac{\delta_2}{\delta_0} + 3\frac{\delta_2^2}{q^2} + 8\frac{\delta_2}{\delta_0}\frac{\delta_2^2}{q^2} - \frac{\delta_2^2}{\delta_0^2}\right)y^2 + \tag{5}$$

$$+ \left(\frac{\delta_2}{\delta_0} - \frac{\delta_2^2}{q^2} - 8\frac{\delta_2}{\delta_0}\frac{\delta_2^2}{q^2} - 4\frac{\delta_2^2}{\delta_0^2}\frac{\delta_2^2}{q^2} + \frac{\delta_2^2}{\delta_0^2}\right)y + 2\frac{\delta_2}{\delta_0}\frac{\delta_2^2}{q^2}\left(1 + \frac{\delta_2}{\delta_0}\right) > 0,$$

$$y < \frac{2}{3}\left(1 + \frac{\delta_2}{\delta_0} - \frac{\delta_2^2}{q^2}\right).$$

We shall call the first stability condition the amplitude condition, and the second the complex condition, while the third will be called the frequency condition.

The stability conditions are best regarded as boundaries which demarcate regions of stable and unstable states on the plane of the resonant curves. The frequency and complex conditions give the equations of the stability boundary if the inequality sign is replaced by the equal sign and δ_2^2/q^2 is eliminated with the aid of the resonant-curve equation (4). We thus obtain the following equation for the stability boundary in the complex condition.

$$\eta^2 = \frac{3(y-1)\left(y-\frac{2}{3}\right)\left[(y-0{,}5)\left(y-\frac{2\delta_2}{\delta_0}\right)+\frac{\delta_2^2}{\delta_0^2}\right]}{(y-0{,}5)\left[-2y(y-1)+8\frac{\delta_2}{\delta_0}(y-0{,}5)-4\frac{\delta_2^2}{\delta_0^2}\right]}\left(\frac{2\delta_2}{\delta_0}-y\right)^2 \tag{6}$$

and in the frequency condition:

$$\eta^2 = \frac{1}{2}\left(y-\frac{2\delta_2}{\delta_0}\right)^3. \tag{7}$$

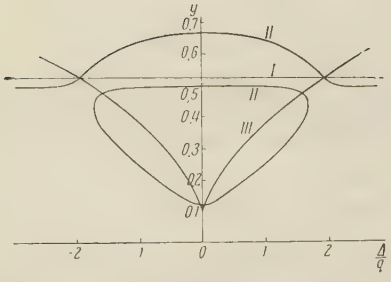


Figure 5. Stability limits for the case $2\delta_2/\delta_0 = 1/4$.

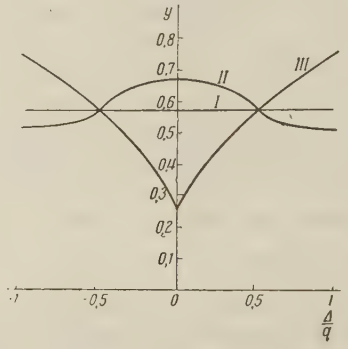


Figure 6. Stability limits for the case $2\delta_2/\delta_0 = 0.1$.

The stability boundaries are shown in Figures 5, 6, and 7. The Roman numbers I, II, and III, denote the boundaries of the amplitude, complex, and frequency stability conditions, respectively.

The frequency condition states that the only stable portions of the amplitude curves are those limited by the points at which the tangent is vertical

$$\frac{dy}{d\left(\nu_1-\frac{\nu_2}{2}\right)} = \infty,$$

and the corresponding external portions of the frequency curves, where

$$\frac{d\omega}{d\left(\nu_1-\frac{\nu_2}{2}\right)} > 0.$$

Equation (7) which we have derived is the equation for the geometric locus of the points at shich the tangent to the amplitude curves is vertical.

The stability boundary (6), corresponding to the complex condition, has for $2\delta_2/\delta_0 < 1/4$ two branches which cross the amplitude curves.

The closed branch lies in the region $2\delta_2/\delta_2 < y < 0{,}5$. When $2\delta_2/\delta_0 = 1/4$ this branch contracts to the point $\Delta = 0$, $y = 3/8$.

The second branch of the boundary (6) always passes through the point $\Delta = 0$, $y = 2/3$ and lies below this point when $2\delta_2/\delta_0 < (1 \div \sqrt{5})/3$. When $2\delta_2/\delta_0 = (1 + \sqrt{5})/3$ this branch turns into the line y = 2/3. When $2\delta_2/\delta_0 > (1 + \sqrt{5})/3$ this branch, passing through the point $\Delta = 0$, y = 2/3 lies above it. When $2\delta_2/\delta_0 > 2$ it closes in the point $\Delta = 0$, $y = 1$.

The stability limit based on the amplitude condition is a straight line which crosses when $2\delta_2/\delta_0 < 2/3$ the boundary of the frequency condition. It is easy to verify that the points of intersection lie on the boundary corresponding to the complex condition and passing through y = 2/3. It is seen therefore that satisfying the complex and frequency conditions is for $2\delta_2/\delta_0 < 2/3$ the necessary and sufficient condition for stability.

As $2\delta_2/\delta_0 \to 2/3$ the points of intersection of all three boundaries come together and merge at $2\delta_2/\delta_0 = 2/3$ in the point $\Delta = 0$, $y = 2/3$. When $2\delta_2/\delta_0 \geqslant 2/3$ the complex condition becomes

the necessary and sufficient stability condition.

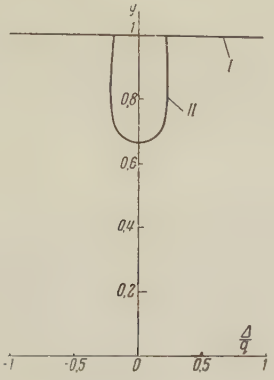


Figure 7. Stability limits for the case $2\delta_2/\delta_0 = 2$.

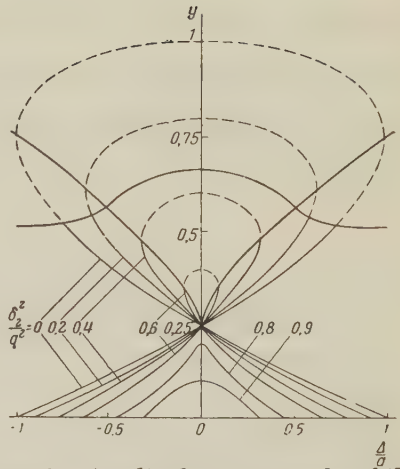


Figure 8. Amplitude curves and stability limits for the case $2\delta_2/\delta_0 = 1/4$. The dashed curves denote unstable states.

Along with determining the stability, it is convenient to consider the behavior of the generator for the case of weak $(2\delta_2/\delta_0 < 1/4)$, medium $(1/4 \leqslant 2\delta_2/\delta_0 < 2/3)$ and strong $(2\delta_2/\delta_0 \geqslant 2/3)$ generator loading.

In the case of medium generator loading (Figure 8) when $1 - \frac{2\delta_2}{\delta_0} \leqslant \frac{\delta_2^2}{q^2} \geqslant \frac{1}{4}\left(1 + \frac{\delta_2}{\delta_0}\right)$ the jump in frequency and amplitude, which takes place, for example, when the detuning of the partial frequencies is changed, occurs at the points where the tangents to the frequency and amplitude curves are vertical. When $\frac{\delta_2^2}{q^2} < \frac{1}{4}\left(1 + \frac{\delta_2}{\delta_0}\right)$ the jump takes place earlier, in view of the complex stability condition.

In the case of weak generator loading (Figure 9) the occurrence of a closed branch of

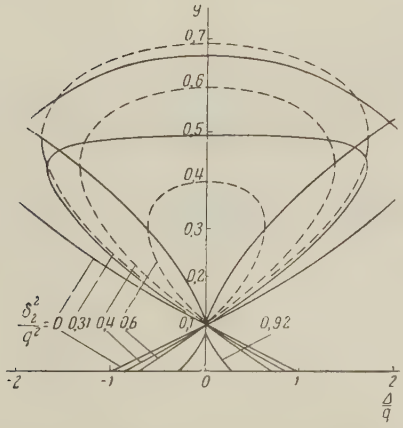


Figure 9. Amplitude curves and stability limits for the case $2\delta_2/\delta_0 = 0.1$. The dashed curves denote unstable states.

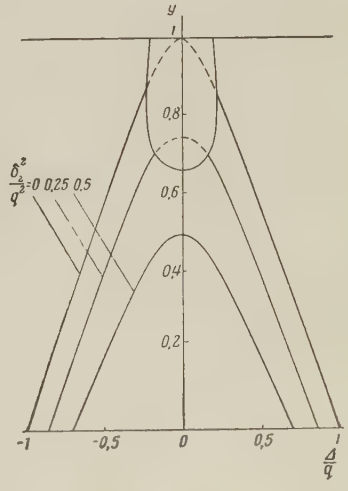


Figure 10. Amplitude curves and stability limits for the case $2\delta_2/\delta_0 = 2$. The dashed curves denote unstable states.

the boundary based on the complex condition makes it possible to reach the points with vertical tangents, when $1-\frac{2\delta_2}{\delta_0} > \frac{\delta_2^2}{q^2} \geqslant \frac{1}{4}\left(1+\frac{\delta_2}{\delta_0}\right)$ only for values δ_2^2/q^2 which are somewhat less than $1 - 2\delta_2/\delta_0$ but which do not much exceed $\frac{1}{4}\left(1+\frac{\delta_2}{\delta_0}\right)$.

In the case of very weak generator loading ($2\delta_2 \to 0$) it is possible to obtain closed frequency-pulling loops on the amplitude curves only when $\delta_2^2/q^2 \sim 0$. For other values of δ_2^2/q^2 the amplitude curves have no stable points above the point of self-intersection of these curves, $y = 2\delta_2/\delta_0$, with the exception of the stable isolated portions of the amplitude curves, appearing when $y \simeq 0.5$, for which $\frac{\delta_2^2}{q^2} \sim \frac{1}{4}\left(1+\frac{\delta_2}{\delta_0}\right)$, and which cannot be obtained by merely changing the detunings. In order to reach these portions, it is necessary to set the necessary detuning with $\delta_2^2/q_2 \sim 0$, and then reduce the coupling coefficient m.

In the case of strong loading (Figure 10) the only single-valued amplitude and frequency curves which can be completely stable are those for which $\delta_2^2/q^2 > 1/3$. When $\delta_2^2/q^2 < 1/3$ there are always unstable portions in the region of near-zero detunings. Oscillation cannot drop here, as occurs in an ordinary oscillator with two degrees of freedom. This is connected with the fact that self-excitation of the generator oscillations takes place in this case up to a certain amplitude, as in the case of the unloaded generator. Starting with this amplitude parametric oscillations occur in the loaded resonant circuit, and these exert a strong reaction on the generator. Such an action leads to a partial or even total suppression of the oscillations in the generator, and consequently, to cessation of parametric excitation of the resonant circuit. The cycle then repeats periodically.

The experimental investigation carried out has shown that in this case the oscillations are of self-modulation character, and turn into intermittent oscillation in the case of zero detuning.

In conclusion it must be noted that the main conclusions of the foregoing theoretical analys were compared with experiment. This comparison has shown that the laws applying to self-oscillating systems with two parametrically coupled degrees of freedom is qualitatively correct.

The authors consider it their pleasant duty to express deep gratitude to K.F. Teodorchik for a fruitful discussion of the results developed in the article.

REFERENCES

1. L.I. Mandel'shtam, Lectures on Oscillations, Izd, An SSSR, 1955.
2. S.M. Rytov, A.M. Prokhorov, M.E. Zhabotinskiy, Zh. Eksp. Teoret. Fiz. 1945, 15, 10, 557 — 571; 11, 613 — 628.
3. Yu.B. Kobzarev, Radiotekhnika, 1950, 5, 2, 41.
4. K.F. Teodorchik, Oscillating Systems, GTTI, 1952.
5. A.A. Vitt, G.S. Gorelik, Zh. Tekh. Fiz. 1933, 3, 2—3, 294.
6. A.N. Charakhch'yan, Zh. Tekh. Fiz., 1936, 6, 7, 1230.
7. Yu. A. Kravtsov, Saturation in Semiconductor-Diode Amplifiers. Abstracts of Papers at the Fourth All-Union Conference of the Ministry of Higher Education USSR on Radio Electronics, Khar'kov, 1960, 68.

Submitted to the editors 31 January 1961

TIME (SPECTRAL) CHARACTERISTICS OF PHASE-DIFFERENCE FLUCTUATIONS OCCURRING IN THE PROPAGATION OF RADIO WAVES IN THE TROPOSPHERE

A.V. Men'

Average "statistical" spectral characteristics of phase-difference fluctuations, obtained from experimental measurements over sea water, of fluctuations on a wavelength $\lambda = 10$ cm and on a fixed path 33 km long, are given for a wide frequency interval.

The data obtained pertain to the case of "transverse" correlation with different distances between receiving antennas (bases) b, ranging from 2 to 100 m, corresponding to different ratios $d/l \gtrless 1$ (l — characteristic correlation scale). The measurements were carried out in the illuminated region as well as in the penumbral and shadow regions.

An experimental investigation of the temporal, and particularly the spectral characteristics of the fluctuations that arise when radio waves propagate in the troposphere is of interest to propagation theory, and also to a study of the physical processes causing these fluctuations. In many earlier theoretical and experimental works devoted to fluctuations [1—10], the principal attention was paid to a determination of their second moments (mean squares, spatial correlation functions), and only recently, in [15], were there reported the results of detailed investigations of the spectral characteristics of the fluctuations of the absolute phase (electric path length),* carried out under conditions close to unbounded space. The present work is devoted to an experimental determination of the spectral characteristics of the fluctuations of the phase differences under conditions where the measurements were carried out directly above the separation boundary (sea). It must be noted, however, that some data, particularly qualitative ones, concerning the spectra of the difference-phase fluctuations have already been cited in the literature [11—14], but in most cases the spectral characteristics were not completely given. Already in this preliminary research it was noted that the low-frequency fluctuation apectra were contained in the interval from hundreds (tens) of cycles down to very low frequencies, on the order of 10^{-3} cps and below. An essential difference is the nonstationary nature of these oscillations, leading to "bursts" and varying spectral distribution with time.

Taking into account the nonstationary nature of the fluctuations in the absolute values of the phases and amplitudes, occurring during propagation in the troposphere, we note that measurement of their temporal or spatial increments (which indeed are the fluctuations of the phase differences in the case of phase measurements) are of special interest. It is known that for nonstationary functions $f(t)$, replacement of these functions by their increments in time $f(t) - f(t_1)$ or in space $f(\vec{r}) - f(\vec{r}_1)$ can lead, for sufficiently small increments, to an appreciable stationarization of the process.

In measurements such as the determination of the spatial phase differences, this interest is due also to the fact that the phase differences have a direct practical application in angle-measuring systems. The expected increased stationarity of these fluctuations, as compared with the fluctuations of the absolute phase, may be connected here with the fact that variations that are slow in time, which in themselves cause the nonstationarity of the processes, should be appreciably correlated at different points in space if the distances between these points are not too great.

One should therefore observe a relative attenuation of the low-frequency spectral components in difference effects.

---

*By absolute phase is meant the phase of the wave at the point of reception, measured relative to the point of radiation.

Naturally, in such measurements it is desirable to obtain the fluctuation spectra in as wide a band as possible, with minimum distortion. We note that whereas the determination of the fluctuation spectrum at high frequencies ($10^2 - 10^1$ cps) entails no special difficulties, the accuracy of determination of very low components is usually limited by the degree of precision of the apparatus and by the finite measurement time.

In the work described here, we investigated the spectrum of fluctuations in the interval from 0.005 or 0.01 to 100 or 200 cps. The experimental measurements were carried out over water and above a fixed path of length greater than 30 km, on a wavelength of 10 cm with vertical polarization of the radiation. The procedure of this research, similar to that reported in [13], consisted of measuring and recording the phase-difference fluctuations ($\delta\varphi_1 - \delta\varphi_d$) between a system of receiving antennas, spaced from 2 to 100 m apart and located along a line perpendicular to the propagation direction*. To increase the measurement accuracy, the experiments were carried out with fixed positions of the receiving and transmitting antennas. The heights of the receiving antennas (h) above sea level were about 4m, while those of the transmitting antennas (h_0) were 35, 18, and 9 m respectively. This ensured reception in both the illuminated region and in the penumbral and shadow regions. The measurements were carried out over a long period of time (in the summer, fall, and winter) under different meteorological conditions, and in each experiment the records were made for all three heights of the transmitter and for distances of 2, 5, 10, 30, and 100 m between receiving antennas.

To estimate the nonstationarity, each measurement, which lasted 5 — 10 minutes, was repeated twice. The spectral characteristics of the fluctuations were determined by different methods by subsequent numerical processing of the resultant records. One of the methods consisted of discretization of the continuous record (determining a series of discrete readings spaced by equal intervals) and determination of the Fourier components at a series of discrete frequencies by numerical means using digital computers with the aid of program tables. These tables were compiled by the "180 ordinate" method in such a way that the processing interval corresponded to 1 — 3 periods of the lowest frequency, while the discretization interval (time between neighboring readings) was not greater than half a cycle of the highest frequency. In this case the frequency interval F_{max}/F_{min} covered by the calculation was not less than 30. If the spectrum had to be determined in a broader frequency range, the calculation was carried out several times with constant total recording duration and different discretization intervals. This made it possible to reduce somewhat the total volume of the calculations, owing to the economy in calculation of the low-frequency components, in which the total number of readings could be appreciably decreased without reducing the accuracy.

Another "continuous" method of processing was based on the use of a mechanical correlator (the amplitudes of the sine and cosine Fourier components were calculated as correlation functions between the investigated process and harmonic oscillations of different frequencies).

All these numerical methods made it possible to determine the running intensity spectra** (the quantity

$$\frac{2}{T^2}\left|\int_T f(t)\, e^{-j\omega t}\, dt\right|^2,$$

where T is the duration of the expansion interval), corresponding physically to the analysis of the process with the aid of a so-called parallel analyzer, consisting of a series of narrow-band filters with analysis band $\Delta F_a \simeq 1/T$, tuned to different frequencies [16,17].

In our investigation, the task was to determine the spectral characteristics of the phase-difference fluctuations, as functions of the height of the transmitter antenna h_0 and of the distance d between receiving antennas (bases) under different meteorological conditions. For this purpose we calculated more than 600 "running" spectra of phase-difference fluctuations, an example of which is Figure 1.

In this figure we show the dependence of the intensity of fluctuations $P'_{\Delta F}$ on the frequency in the interval from 0.01 to 10 cps, at an equivalent analysis band approximately 0.006 cps, normalized in each case to the total fluctuation intensity, as given by the value of the

---

*Such a placement of the antennas corresponds to measuring the so-called transverse spatial correlation of the fluctuations.

**That is, a spectral expansion (of the function) corresponding to the given time interval.

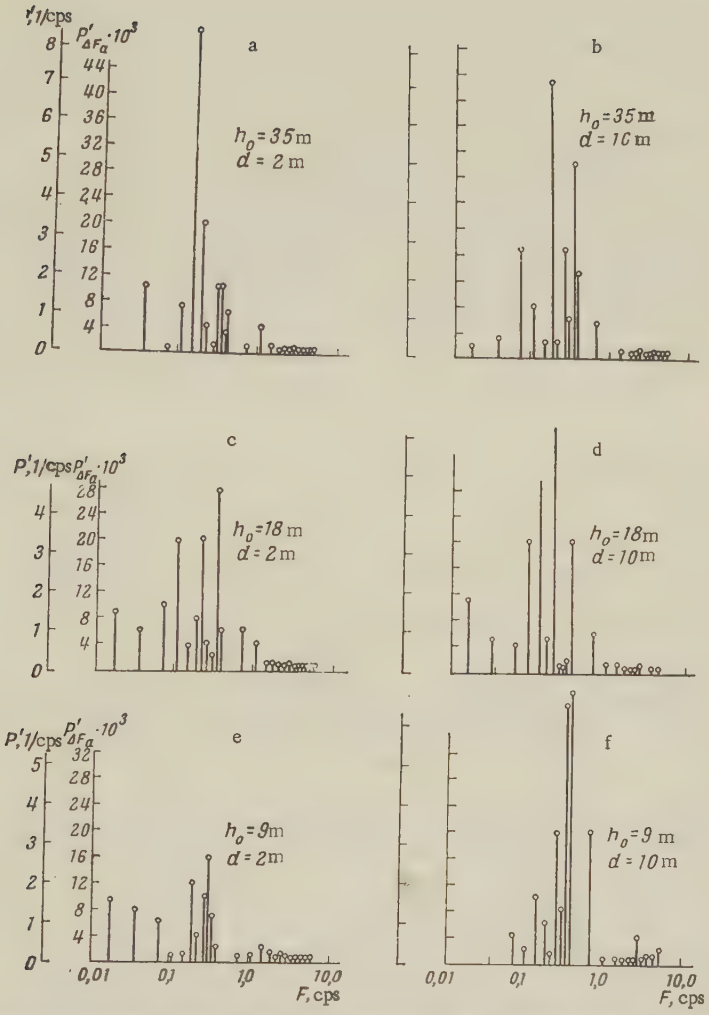


Figure 1. Dimensionless running spectra of intensity of fluctuations of phase difference, P, $P'_{\Delta Fa} = f(F)$; $P'_{\Delta Fa}$ — intensity of fluctuations in band $\Delta Fa = 0.006$ cps, normalized to the total fluctuation intensity; F — fluctuation frequency.

mean squares. The same figure shows the scale of the quantity p' (in cps^{-1}), defined as $P'_{\Delta Fa}/\Delta Fa$, which can be defined (apart from a constant) as the averaged normalized spectral density of the intensity, a density satisfying the normalization condition $\int_F p'dF = 1$. As follows from Figure 1. qualitatively similar running spectra are noted in this experiment, independently of the height of the transmitter h_0 and the distance d between receiving antennas. We note that the total intensity of the fluctuations was in these measurements also practically independent of the height h_0. A typical appreciable inhomogeneity in the distribution of the intensity over the frequencies (bursts) was observed, and this made the analysis of the main laws quite difficult.

In individual cases, for example in analysis within the limits of an interval of about 0.006 cps, which constitutes not more than one percent of the total frequency band, bursts reaching 10 — 15 percent of the total fluctuation intensity were observed, indicating an increased content

of individual harmonic components in the spectrum .

It follows from the foregoing data that at higher frequencies of the spectrum there is usually a tendency for a sharp decrease in the spectral density with increasing frequency, a tendency most clearly pronounced at maximum values of d and h_0. It is typical of this experiment that the greatest spectral densities correspond to a frequency interval from 0.1 to 1 cps, although in most measurements made in this interval there was usually also an increase in fluctuation intensity with decreasing F.

An analysis of a large number of measurements of the running fluctuation spectra has confirmed the results of the preliminary measurements [14] with respect to their essential nonstationarity. Whereas the total intensity of the fluctuations was in many experiments relatively stable with accuracy to 0.5 — 1 db (quasi-stationary measurements) and only in individual cases did the changes reach 3 — 5 db (nonstationary measurements), at the same time within narrow frequency intervals, as a rule, considerably greater changes in the intensity were observed, on the order of 10 — 20 db and more. The nonstationarity of the running spectra was also manifest in their dependence on the duration of the measurement interval. As the measurement durations were varied one observed along with a

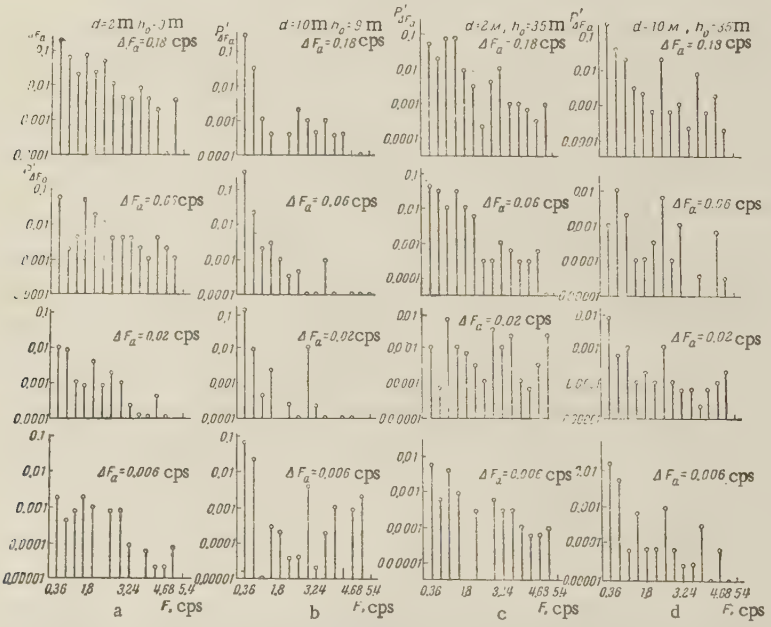


Figure 2. Dependence of the running spectra of the fluctuations on the duration of the measurement interval, determining the equivalent analysis band ΔF.

certain change in the total fluctuation intensity (usually the intensity increased with increasing measurement interval, owing to the appearance of slower and slower components), also an appreciable redistribution of intensity over the frequencies.

Typical measurements of running spectra are illustrated in Figure 2 as functions of the measurement-interval duration (processing-interval). This figure shows the values of the fluctuation power contained in the band ΔF_a, calculated from the experimental recordings, at different frequencies in the interval from 0.36 to 5.04 cps at equivalent analysis band values (ΔF_a) of 0.18, 0.06, 0.02, and 0.006 cps. An analysis of such calculations shows that on the average there is a tendency toward reduction of intensity of fluctuations contained in a definite band, as the band narrows down, although in this case there is a great spread in the calculated data, owing to the nonstationarity.

By way of example, we can point to Figure 3, which shows for several settings of F from 0.36 to 2.16 cps, the variation with the duration of the analysis T of the intensity

of fluctuations P_T, contained in a band approximately 1/T, referred to the value P_{T_0} corresponding to the minimum time $T_0 = 5.6$ sec. As follows from Figure 3, for all the settings given, the reduction in power with increasing analysis time agrees in the mean, albeit with a large spread, with the proportionality to T^{-1} inherent in stationary processes of the type of continuous "white" noise.

It is also advantageous to determine the dependence of the average spectral density, (defined as the ratio of the intensity of fluctuations $P'_{\Delta Fa}$ to the bandwidth ΔFa in which there are included) on the frequency and on the analysis band. Similar results for distances d between receiving antennas of 2 and 10 m and heights h_0 of transmitter antennas 9, 8, and 35 m are shown in Figure 4, where different symbols are used for data corresponding to analysis with a band from 0.006 to 0.18 cps. The dashed curves show the averaged functions p'(F), relative to which the spread of the experimental data was on the average ± 5 db, increasing in some cases to ± 10 db and more. As follows from Figure 4, in spite of the

Figure 3. Variation of fluctuation intensity P_T included in the band 1/T, as a function of the analysis time and of the tuned frequency f, d = 2 m, h_0 = 18 m:
1 — F = 0.36 cps 2 — F = 0.72 cps 3 — F = 1.08 cps 4 — F = 1.44 cps 5 — F = 1.8 cps 6 — F = 29.1 cps

spread in the data, a rapid decrease in spectral density with increasing frequency was noticed in all cases. The averaged function p'(F) was approximately proportional here to $F^{-1.5}$ for a base d of 2 m at the reception point, and proportional to $F^{-2.5}$ for d = 10 m. At the same time, the effect of the height on p'(F) was less clearly pronounced in this case.

One can note as a result of a large number of measurements that in general the altitude dependence of the spectral characteristics of the phase-difference fluctuations, as well as the altitude characteristic of the over-all fluctuation intensity [13, 14] is subject to considerable changes from experiment to experiment, and deviations from typical standard conditions are observed here much more frequently than, for example, in the dependence of the spectral characteristics on the size of the base. The data shown in Figure 4, like the data in the preceding figures, pertain precisely to one such experiment, where an anomaly was observed in the distribution of the over-all intensity of the fluctuations with height. Unlike the ordinary altitude dependence of the fluctuations $\overline{(\delta\varphi_1 - \delta\varphi_d)^2} \sim h_0^{-2}$ [8, 14], in this case it was not seen in explicit form. It must be noted that the dependence of the spectral density on the frequency, noted in Figure 4, which for d = 10 m is on the average proportional to $F^{-2.5}$, is in good agreement with the results of investigation of fluctuations of the absolute phase $\delta\varphi$ [15], where an analogous characteristic of the spectral density over a wide frequency interval was approximated by the law $F^{-2.8}$.

Thus, at sufficiently large bases $d(d \gg l)$, where l is the characteristic dimension necessary to obtain an appreciable decrease in the correlation of the fluctuations of the dielectric constant of a turbulent inhomogeneous medium [2 —5]) the spectral densities of the fluctuations of the absolute phase and of the phase difference practically coincide, apart from a constant, particularly in the region of the high-frequency components. It can be shown that this result is not accidental. Representing the autocorrelation function of the phase-difference fluctuations B(t) in the form

$$B(\tau) = \overline{[\delta\varphi_1(t) - \delta\varphi_d(t)][\delta\varphi_1(t+\tau) - \delta\varphi_d(t+\tau)]} \qquad (1)$$

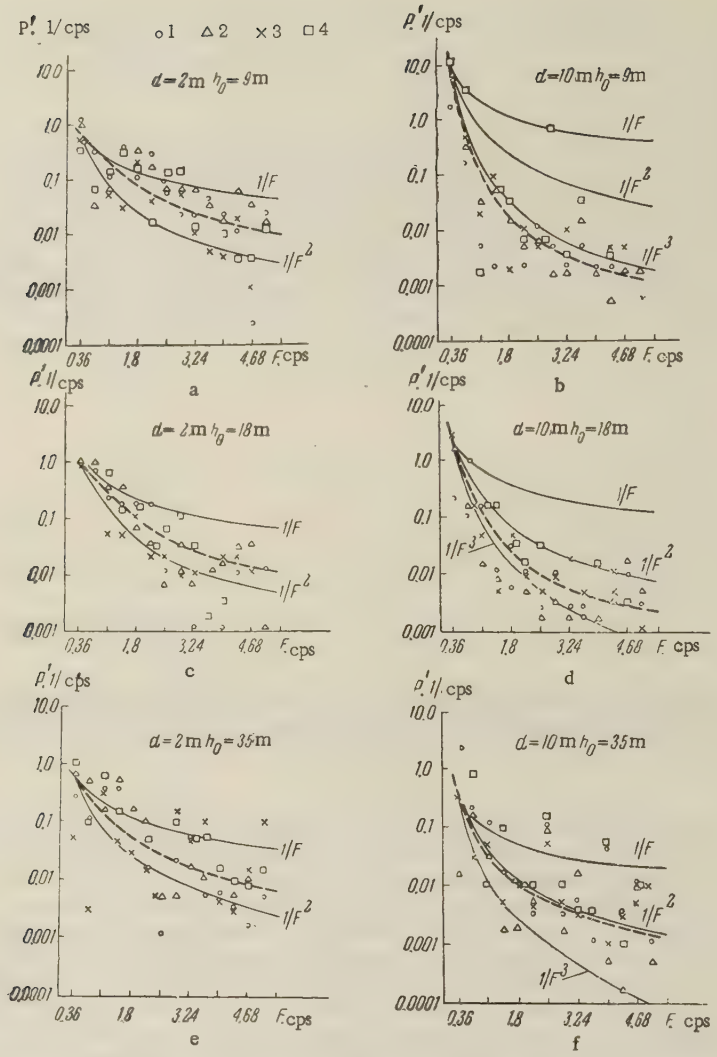


Figure 4. Dependence of the averaged spectral density of the phase-difference fluctuations p' (for analysis with different band F_a) on the frequency F, on the base d, and on the height h0:

1 — $\Delta F_a = 0.18$ cps 2 — $\Delta F_a = 0.06$ cps 3 — $\Delta F_a = 0.02$ cps 4 — $\Delta F_a = 0.006$ cps

and recognizing that when $d \gg l$, except for a narrow sector of angles along the propagation direction [18], the spatial correlation terms of the type $\overline{\delta\varphi_1 \delta\varphi_d}$ tend to zero if averaged over a sufficient interval, we get

$$B(\tau) \simeq B_1(\tau) + B_d(\tau); \tag{2}$$

Defining the spectral density as the Fourier transform of the correlation function B(t),

$$p(\omega) = p_1(\omega) + p_d^{\blacksquare}(\omega) \tag{3}$$

or in averaging over an ensemble of independent measurements, assuming the medium to be homogeneous in the mean (statistically) when $d \ll L$ (L — length of the path), we get

$$\overline{B}_1 = \overline{B}_d, \quad \overline{p}(\omega) = 2\overline{p}_1(\omega). \tag{4}$$

Here $\delta\varphi_1$ and $\delta\varphi_d$ are the fluctuations of the absolute phase at points a distance d apart; $B_1(d)$ (τ) and $p_1(d)$ (ω) are the "running" autocorrelation functions and spectral densities of the absolute-phase fluctuations at the corresponding points.

We note that this case takes place in Figure 4 for a base d = 10 m, since the parameter l measured experimentally on the spatial structure functions, amounted in these measurements to 2 —3 m for all heights.* The reduction in the base entails an increase in the correlation predominantly in the low-frequency components; consequently, their contribution to the intensity of the phase-difference fluctuations increases. Inasmuch as this slows down the increase in the spectral density with decreasing frequency, the results for d = 2 m $(d \simeq l)$ appears natural, the frequency dependence of the spectral density p(F) being less sharply pronounced, an and being approximately proportional to $F^{-1.5}$. Apparently, at large bases one should expect a further weakening of this dependence, that is, a relative broadening of the spectrum of the fluctuations. The nonstationarity of the intensity of the investigated fluctuations, noted above, and the nonstationarity of their running spectra, which is sharply pronounced in the case of finite measurement times, both being apparently due to the presence of components of rather low frequency in the spectra, make it practically impossible in may cases to carry out a direct analysis of the main laws governing the spectral characteristics. To investigate these laws one can use statistical reduction of the running spectra, determined in a series of independent measurements. In such a statistical spectrum the intensity at each setting can be determined as the average over the ensemble, or as values (levels) of intensity exceeded by the experimental data with definite probability. In the latter case one determines for the fluctuation intensity at different frequencies, measured in a definite analysis band ΔF_a, integral distributions such as the probability of the intensity exceeding different values.

The result of such a processing of thirty running spectra with $\Delta F_a = 0.18$ cps in the interval 0.36 — 5.04 cps is shown in Figure 5 for different h_0 and d. Special marks on the figure indicate levels with probability W = 50 percent, while the dashed curves represent the averaged function $P'_{\Delta F_a}(F)$ drawn through these points. The same figures show, designated by different symbols, for comparison, the values of the intensity in the indicated band $P'_{\Delta F_a}$ for several settings, referred to the total intensity; these values were obtained in the first six measurements (the interval between measurements was about 5.6 sec).

As follows from the data given, the spread in the values with 50 percent probability relative to the averaged dependence did not exceed in the mean ± 2.5 db, whereas for individual measurements it reached ± 10 db and more. For the values averaged with 50 percent probability, the intensity (spectral density) varies with the frequency, as $F^{-2.5}$ for d + 2 and as F^{-2} to $F^{-2.5}$ for d = 10 m, as above.

An analogous statistical averaging of more than 600 running spectra, determined in a lower-frequency interval 0.03 — 0.42 cps at a band $\Delta F_a = 0.006$ cps, is illustrated in Figure 6. Unlike the data of Figure 5, corresponding to measurements carried out during a single experiment under practically invariant conditions, in this case we averaged the results of investigations carried out over a long period of time under a great variety of meteorological conditions.

The correlation in space in such a synthesis of a fluctuation spectrum corresponded to distances l = 10 — 30 m. As follows from Figure 6, it is possible in all cases to draw through the statistically determined values of spectral density with W = 50 percent, with relatively small spread, averaged functions p'(F) which range (depending on the base) from

---

* An experimental investigation of spatial structure functions $\overline{(\delta\varphi_1 - \delta\varphi_d)^2} = f(d)$ is described in [13, 14]. It is shown there that the fluctuations due only to low-frequency spectral components are correlated in space over greater distances than fluctuations due only to high frequency spectral components.

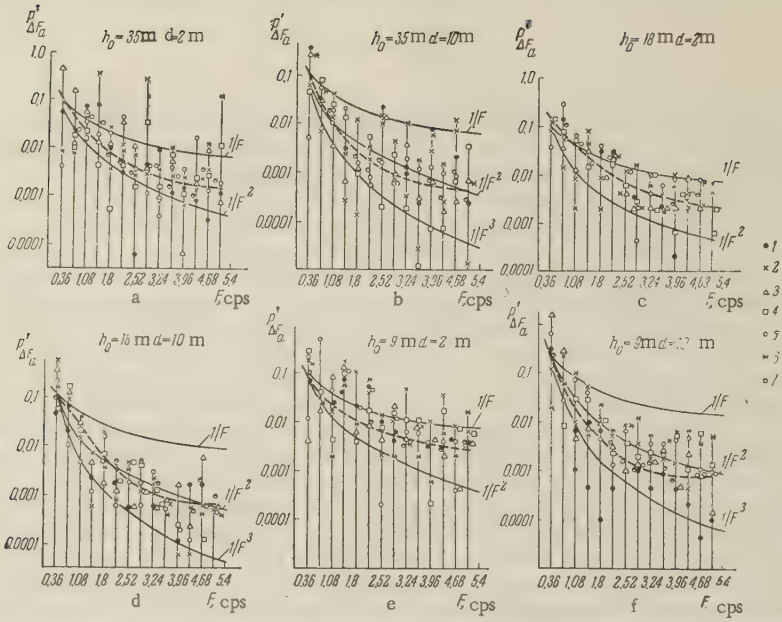


Figure 5. Comparison of the running fluctuation spectra with the averaged dependence of the corresponding probability of exceeding W = 50 percent (dashed) at different h_0 and d.

1 — first measurement; 2 — second measurement, 3 — third measurement;
4 — fourth measurement; 5 — fifth measurement; 6 — sixth measurement;
7 — values with W = 50 percent.

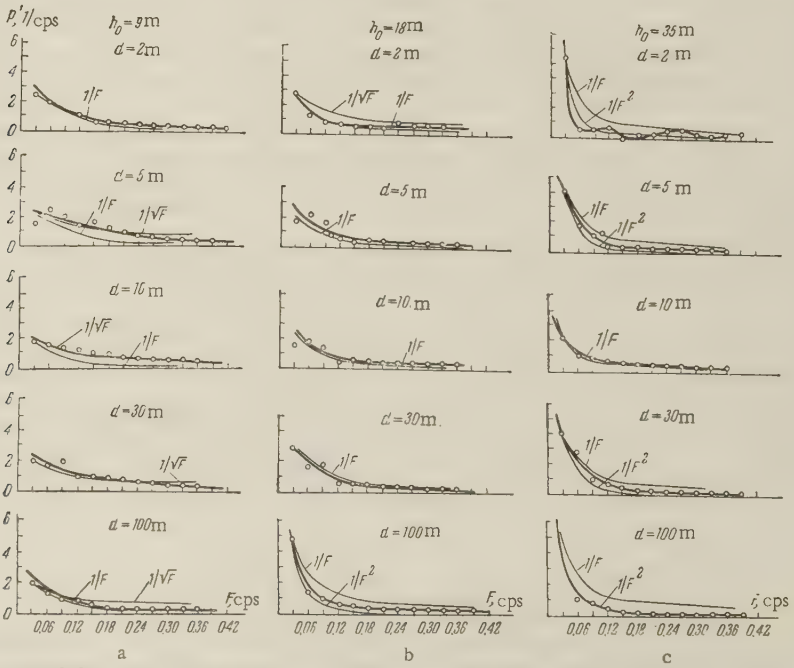


Figure 6. Statistical frequency dependences of the spectral density at different h_0 and d.

$p \sim F^{-0.5} - F^{-1}$ for $h_0 = 9$ m to $p' \sim F^{-1.5} - F^{-2}$ for $h_0 = 35$ m. It is typical that on all plots of Figure 6, down to frequencies on the order of 0.01 cps, the spectral density continues to increase with decreasing frequency although slower than in the high-frequency interval (Figure 5). Thus, statistical processing of the spectra of nonstationary fluctuations makes it possible to determine for different d and h_0 the main tendencies of the spectral characteristics p(F). It must be noted, however, that in many cases it is difficult to use this method, for example, in the analysis of the causes of changes in the temporal characteristics of fluctuations, to which, generally speaking, considerable interest is attached. Taking account of the fact that the running spectra are not constant from experiment to experiment, it is deemed advantageous to determine in these experiments the connection between these running spectra and the changes of the meteorological conditions along the path. For this purpose we regularly measured the pressure, temperature, wind, humidity, etc. at four separated points, two of which were at the ends of the path.

In particular, an attempt was made to determine the dependence of the spectral characteristics on the direction and force of the wind, for in the framework of the widely known "frozen turbulence" hypothesis [1, 6] the temporal (spectral) dependences were due only to the transport of the inhomogeneities by the average wind, and in this case the spectra should be appreciably affected only by the "transverse" component of the wind, directed perpendicular to the radial path.

As a result of the comparison made we can note that we were practically unable to observe an explicit connection between the fluctuation spectra and the averaged meteorological conditions (in particular, with the wind direction and force). In the determination of the running spectra, such a comparison, however, could be made only qualitatively, which along with the complexity and laboriousness of the procedure turned out to greatly limit its possibilities. It may prove more advantageous in such cases to make numerical estimates of the spectral fluctuation characteristics, obtained approximately, without resorting to a detailed investigation of the frequency dependence of the spectral density. This question however, is beyond the scope of the present article.

REFERENCES

1. V.A. Krasil'nikov, Izvestiya Akad. Nauk SSSR, Geographic and Geophysics Series, 1949, 13, 1, 33.
2. A.M. Obukhov, Izvestiya Akad Nauk SSSR, Geophysics Series 1953, 17, 2, 155.
3. L.A. Chernov, Doklady AN SSSR, 1954, 98, 6, 953; Akust. Zh. 1957, 3, 2, 192.
4. L.A. Chernov, Propagation of Waves in Medium with Random Inhomogeneities, Izd. AN SSSR, 1958.
5. V.N. Karavaynikov, Akust. Zh. 1958, 3, 2, 165.
6. R.B. Muchmore, A.D. Wheelon, Proc. I.R.E., 1955, 43, 10, 1437.
7. A.V. Men', S. Ya. Braude, V.I. Gorbach, Doklady, Akad. Nauk, Ukrainian SSR, 1959, 3, 7, 740.
8. A.V. Men', V.I. Gorbach, S.Ya. Braude, Izv. Vuzov, MVR SSSR, — Radiophysics, 1959, 2, 3, 388.
9. E.A. Kaner, and F.G. Bass, Doklady AN SSSR, 1959, 127, 4, 792.
10. E.A. Kaner, F.G. Bass, Izv. Vuzov MVR SSST, Radiophysics, 1959, 2, 4, 553.
11. A.P. Deam, B.M. Fannin, Proc. I.R.E., 1955, 43, 10, 1402.
12. J. W. Herbstreit, M.C. Thompson, Proc, I.R.E., 1955, 43, 10, 1391.
13. A.V. Men', S.Ya. Braude, V.I. Gorbach, Doklady AN SSSR, 1959, 125, 5, 1019.
14. A.V. Men', S. Ya. Braude, V.I. Gorbach, Izv. Vuzov MVO SSSR, Radiophysics, 1959, 2, 6, 848.
15. M.C. Thompson, H.B. Janes, J. Res. Nat. Bur. Standards, 1959, 63D, 4, 45.
16. A.A. Kharkevich, Spectra and Analysis, GTTI, 1953,
17. V.I. Bunimovich, Fluctuation Processes in Radio Receivers, Soviet Radio Press, 1951.
18. A.V. Men', Izv. Vuzov, MVR SSSR, Radiophysics, 1959, 2, 3, 395.

Submitted to the editors 16 November 1960

FREQUENCY CORRELATION OF AMPLITUDE AND PHASE OF FLUCTUATIONS IN THE USE OF A SHARPLY DIRECTIONAL ANTENNA

M.F. Bakhareva

The finite dimensions of the receiving equipment were taken into account in a calculation of the coefficients of frequency correlation of the amplitude fluctuations R_A^Σ and phase fluctuations R_φ^Σ of the electromagnetic field of an originally plane wave that has passed through a layer of thickness L in a medium with random large-scale refractive index inhomogeneities. Plots of R_A^Σ and of R_φ^Σ are prepared as functions of the frequency shift δ, the wave parameter $D = 4L/ka^2$ (where L is the thickness of the layer, k the wave number, and a the mean dimension of the inhomogeneities of the medium), and of the relative dimension h/a of the receiver antenna.

INTRODUCTION

The calculation of correlation functions of amplitude and phase fluctuations of a plane wave passing through a statistically inhomogeneous layer has been the subject of several investigations [1 — 3].

In [3] the coefficient of frequency correlation of the amplitude and phase fluctuations, R_A and R_φ, was determined as a function of D and δ at a point, that is, for infinitesimally small receiver dimensions.

In the present paper we derive expressions for the coefficients R_A^Σ and R_A^Σ of the frequency correlation of amplitude and phase fluctuations in the path of a sharply directional antenna (lens) as functions of D, δ, and the relative dimension of the antenna. We first calculate the frequency-transverse correlation functions of the amplitude and phase fluctuations, that is, the correlations at two points perpendicular to the propagation, and accordingly, for two different frequencies, since these functions of the determine the frequency correlation at the focus of the lens. In the derivation of the formulas for R_A^Σ and R_φ^Σ it is assumed that the field fluctuations are small and that the antenna has the form of a square with side h.

In the particular case when h/a = 0, the formulas obtained become identical with those previously obtained in [3].

1. FORMULATION OF THE PROBLEM

Assuming two plane monochromatic waves of frequencies ω_1 and ω_2 to pass in the ox direction through a statistically inhomogeneous layer of thickness L, and assuming the lens to be located in the plane x + L (Figure 1), let us calculate the coefficients of the frequency correlation of the amplitude fluctuations R_A^Σ and of the phase fluctuations R_φ^Σ at the focus of the lens, namely,

$$R_A^\Sigma = \frac{\overline{\Delta A_1^\Sigma \Delta A_2^\Sigma}}{\sqrt{\overline{(\Delta A_1^\Sigma)^2}\ \overline{(\Delta A_2^\Sigma)^2}}}, \qquad R_\varphi^\Sigma = \frac{\overline{\Delta \varphi_1^\Sigma\ \Delta \varphi_2^\Sigma}}{\sqrt{\overline{(\Delta \varphi_1^\Sigma)^2}\,\overline{(\Delta \varphi_2^\Sigma)^2}}}, \tag{1}$$

where ΔA_i^{Σ} and $\Delta\varphi_i^{\Sigma}$ are respectively the fluctuations of the amplitude and of the phase at the focus of the lens, at frequency ω_i

We shall calculate R_A^{Σ} and R_{φ}^{Σ} for small perturbations of the fields of the both waves. In this case, as shown in [4], we can assume that the fluctuations of the amplitude and the phase at the focus are due respectively to the fluctuations of the amplitude and phase in the wave incident on the lens. On this bases, we can write for ΔA_i^{Σ} and $\Delta\varphi_i^{\Sigma}$

$$\Delta A_i^{\Sigma} = \frac{1}{\lambda_i F}\int_{\Sigma} \Delta A_i(s)\,ds = \frac{\overline{A}}{A_0}\frac{1}{\Sigma}\int_{\Sigma}\Delta A_i(s)\,ds; \quad \Delta\varphi_i^{\Sigma} = \frac{1}{\Sigma}\int_{\Sigma}\Delta\varphi_i(s)\,ds, \tag{2}$$

where F is the focal length; $\overline{A}$ the mean amplitude of the field at the focus; A_0 the amplitude of the incident wave; Σ the area of lens: $\Delta A_i(s)$ $[\Delta\varphi_i(s)]$ the fluctuation of the amplitude (phase) at the point s at frequency $\omega_i (i = 1, 2)$.

Substituting (2) in (1) and averaging, we obtain

$$R_A^{\Sigma} = \frac{\iint_{\Sigma\Sigma'} \overline{\Delta A_1(s)\,\Delta A_2(s')}\,ds\,ds'}{\sqrt{\iint_{\Sigma\Sigma'}\overline{\Delta A_1(s)\,\Delta A_1(s')}\,ds\,ds' \iint_{\Sigma\Sigma'}\overline{\Delta A_2(s)\,\Delta A_2(s')}\,ds\,ds'}},$$

$$R_{\varphi}^{\Sigma} = \frac{\iint_{\Sigma\Sigma'} \overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s)}\,ds\,ds'}{\sqrt{\iint_{\Sigma\Sigma'}\overline{\Delta\varphi_1(s)\,\Delta\varphi_1(s')}\,ds\,ds' \iint_{\Sigma\Sigma'}\overline{\Delta\varphi_2(s)\,\Delta\varphi_2(s')}\,ds\,ds'}}. \tag{3}$$

The transverse correlations $\overline{\Delta A_i(s)\,\Delta A_i(s')}$ and $\overline{\Delta\varphi_i(s)\Delta\varphi_i(s')}$ $(i = 1, 2)$, contained in the integrand of the denominator of (3), have been evaluated in [1]. The integrals of the numerator of (3) also contain the functions $\overline{\Delta A_1(s)\,\Delta A_2(s')}$ and $\overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s')}$ which we shall call the frequency-transverse correlation functions of the fluctuations of the amplitudes and phases, respectively. The next section is devoted to a calculation of these functions.

2. FREQUENCY-TRANSVERSE CORRELATION FUNCTIONS OF THE AMPLITUDE AND PHASE FLUCTUATIONS

Using the results of [5] [Chapter V, formulas (72) and (71)] and the notation of Figure 1, let us write down expressions for small amplitude and phase fluctuations at two points s and s' with coordinates (L, 0, 0) and (L, l, 0) in the form of integrals over a scattering layer with volume V:

$$\left.\begin{matrix}\frac{\Delta A_i(s)}{A_0}\\ \Delta\varphi_i(s)\end{matrix}\right\} = \frac{k_i^2}{2\pi}\int_V \left.\begin{matrix}\cos\\ \sin\end{matrix}\right\} k_i[r - (L-\xi)]\frac{\mu(\xi,\eta,\zeta)}{r}\,d\xi\,d\eta\,d\zeta,$$

$$\left.\begin{matrix}\frac{\Delta A_i(s')}{A_0}\\ \Delta\varphi_i(s')\end{matrix}\right\} = \frac{k_i^2}{2\pi}\int_V \left.\begin{matrix}\cos\\ \sin\end{matrix}\right\} k_i[r_l - (L-\xi)]\frac{\mu(\xi,\eta,\zeta)}{r_l}\,d\xi\,d\eta\,d\zeta, \tag{4}$$

where A_0 is the amplitude of the waves incident on the layer (the same for both frequencies; k_i is the wave number at frequency ω_i $(i = 1, 2)$; $\mu(\xi, \eta, \zeta)$ is the fluctuation of the refractive index at the point (ξ, η, ζ); $r = \sqrt{(L-\xi)^2 + \eta^2 + \zeta^2}$ $r_i = \sqrt{(L-\xi)^2 + (l-\eta 2) + \zeta^2}$ are the distances between points s, s', and the point with coordinates (ξ, η, ζ) inside the scattering volume.

Multiplying the values of $\frac{\Delta A_i}{A_0}(\Delta\varphi_i)$ for different frequencies and at different points,

averaging, and performing calculations similar to those used for the transverse correlation of amplitude and phase fluctuations [1], we obtain on the basis of (4)

$$\left.\begin{matrix}\overline{\frac{\Delta A_1(s)\,\Delta A_2(s')}{A_0^2}}\\ \overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s')}\end{matrix}\right\} = \frac{1}{2}\overline{\mu^2}k_1k_2\,(J_1 \mp J_2), \tag{5}$$

where

$$J_{1,2} = \frac{1}{2j}\int_0^L dx \int_{-\infty}^{\infty} \exp\left(-\frac{\xi^2}{a^2}\right)\left\{\frac{\exp\left[\frac{j\frac{l^2}{a^2}}{b_{1,2}(x,\,\xi)-j}\right]}{b_{1,2}(x,\,\xi)-j} - \frac{\exp\left[-\frac{j\frac{l^2}{a^2}}{b_{1,2}(x,\,\xi)+j}\right]}{b_{1,2}(x,\,\xi)+j}\right\} d\xi \tag{6}$$

and where we used the symbols

$$\begin{aligned} b_1(x,\,\xi) &= \frac{2(L-x)}{a^2}\left(\frac{1}{k_1}-\frac{1}{k_2}\right) - \frac{\xi}{a^2}\left(\frac{1}{k_1}+\frac{1}{k_2}\right),\\ b_2(x,\,\xi) &= \frac{2(L-x)}{a^2}\left(\frac{1}{k_1}+\frac{1}{k_2}\right) - \frac{\xi}{a^2}\left(\frac{1}{k_1}-\frac{1}{k_2}\right),\\ j &= \sqrt{-1}. \end{aligned} \tag{7}$$

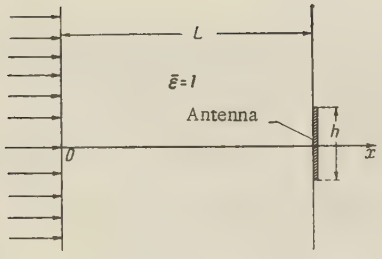


Figure 1. Diagram showing location of receiving antenna.

In the derivation of (5) we have assumed that the fluctuations of the refractive index have a normal distribution, that is,

$$\overline{\mu(\xi,\,\eta,\,\zeta)\,\mu(\xi',\,\eta',\,\zeta')} = \overline{\mu^2}\exp\left[-\frac{\sqrt{(\xi-\xi')^2+(\eta-\eta')^2+(\zeta-\zeta')^2}}{a^2}\right],$$

where $\overline{\mu^2}$ and a are, respectively, the intensity and mean dimensions of these fluctuations; $k_i a \gg 1\,(i = 1, 2)$ and $L \gg a$.

Making the substitution $z = \frac{j(l^2/a^2)}{b_{1,2}(x,\xi)-j}$ in the first integrand of (6) and the substitution $z = \frac{-j\,(l^2/a^2)}{b_{1,2}(x,\xi)+j}$ in the second, we obtain after integrating with respect to x

$$J_{1,2} = \frac{a^2}{4j\left(\frac{1}{k_1}\mp\frac{1}{k^2}\right)}\int_{-\infty}^{\infty}\exp\left(-\frac{\xi^2}{a^2}\right)\left\{\mathrm{Ei}\left[\frac{j\frac{l^2}{a^2}}{b_{1,2}(L,\,\xi)-j}\right] - \mathrm{Ei}\left[-\frac{j\frac{l^2}{a^2}}{b_{1,2}(L,\,\xi)+j}\right] - \mathrm{Ei}\left[\frac{j\frac{l^2}{a^2}}{b_{1,2}(0,\,\xi)-j}\right] + \mathrm{Ei}\left[-\frac{j\frac{l^2}{a^2}}{b_{1,2}(0,\,\xi)+j}\right]\right\} d\xi, \tag{8}$$

where E_i is the exponential integral function.

Let us set

$$ {}^1\varepsilon_{1,2} = \frac{i\frac{l^2}{a^2}}{b_{1,2}(L,\xi)-i}, \qquad {}^2\varepsilon_{1,2} = \frac{i\frac{l^2}{a^2}}{b_{1,2}(L,\xi)+i}. $$

According to (7), if $k_1 a \gg 1$, the inequalities $|b_{1,2}(L,\xi)| \ll 1$, are satisfied. We therefore expand the differences $\{E_i({}^1\epsilon_{1,2}) - E_i({}^2\epsilon_{1,2})\}$ under the integral sign in (8) in powers of the small increments of the argument

$$ \Delta\varepsilon_{1,2} = {}^2\varepsilon_{1,2} - {}^1\varepsilon_{1,2} = -\frac{2i\frac{l^2}{a^2}\,b_{1,2}(L,\xi)}{1+b_{1,2}^2(L,\xi)}. $$

As a result we obtain, accurate to $b_{1,2}^3$,

$$ \mathrm{Ei}({}^2\varepsilon_{1,2}) - \mathrm{Ei}({}^1\varepsilon_{1,2}) = 2ib_{1,2}(L,\xi)\exp\left(-\frac{l^2}{a^2}\right)\times \\ \times\left[1-\frac{b_{1,2}^2(L,\xi)}{3}\right]\left(1-2\frac{l^2}{a^2}+\frac{1}{2}\frac{l^4}{a^4}\right)\Big]. \tag{9} $$

Substituting this expansion into (8), discarding the integrals of the odd functions of ξ, and using

$$ k_1 = \frac{\omega}{c}(1-\delta),\ k_2 = \frac{\omega}{c}(1+\delta),\ D = \frac{4cL}{\omega a^2}, $$

we find

$$ J_1 = \frac{a^2\omega(1-\delta^2)}{8ic\delta}\int_{-\infty}^{\infty}\exp\left(-\frac{\xi^2}{a^2}\right)\left\{\mathrm{Ei}\left[-\frac{i\frac{l^2}{a^2}}{b_1(0,\xi)+i}\right]- \right. \\ \left. -\mathrm{Ei}\left[\frac{i\frac{l^2}{a^2}}{b_1(0,\xi)-i}\right]\right\}d\xi, \tag{10} $$

$$ J_2' = \frac{\sqrt{\pi}a^2\omega(1-\delta^2)}{8ic}\left[\mathrm{Ei}\left(-\frac{i\frac{l^2}{a^2}}{\frac{D}{1-\delta^2}+i}\right)-\mathrm{Ei}\left(\frac{i\frac{l^2}{a^2}}{\frac{D}{1-\delta^2}-i}\right)\right]. \tag{11} $$

In the derivation of (11) we have neglected the second term in the expression for $b_2(0,\xi)$ because $k_1 a \gg 1$ and ξ does not exceed a in practice, owing to the rapid falloff of the function $\exp(-\xi^2/a^2)$ under the integral sign in (8).

We carry out the integration in (10) for two particular cases: 1) $D\delta/(1-\delta^2) \ll 1$ and 2) $D\delta/(1-\delta^2) \gg 1$. In order not to violate the requirement that the field fluctuations be small in the second case, we assume that the intensity of the fluctuations of the refractive index $\overline{\mu^2}$ is so small, that the field perturbations remain small even at distances determined by the requirement $D\delta/(1-\delta^2) \gg 1$.

1) $D\delta/(1-\delta^2) \ll 1$. In this case on the bases of 7 the following inequality is satisfied if $k_1 a \gg 1$.

$$ b_1(0,\xi) = \frac{D\delta}{1-\delta^2}\left(1-\frac{\xi}{2L\delta}\right) = \frac{1}{1-\delta^2}\left(D\delta-\frac{2c\xi}{\omega a^2}\right) \ll 1. $$

We can therefore use for the difference in the exponential integral functions in (10) an expansion similar to (9). After integrating (10) with respect to ξ we get

$$J_1 = \sqrt{\pi} aL \exp\left(-\frac{l^2}{a^2}\right)\left\{1-\left(1-2\frac{l^2}{a^2}+\frac{1}{2}\frac{l^4}{a^4}\right)\times\right.$$
$$\left.\times\left[\frac{1}{3}\frac{D^2\delta^2}{(1-\delta^2)^2}+\frac{2c^2}{\omega^2 a^2}\frac{1}{(1-\delta^2)^2}\right]\right\}. \qquad (12)$$

Substituting (11) and (12) into (5) we obtain the following expressions for the frequency-transverse correlation functions of the amplitude and phase fluctuations:

$$\left.\begin{matrix}\overline{\frac{\Delta A_1(s)\,\Delta A_2(s')}{A_0^2}}\\ \overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s')}\end{matrix}\right\} = \frac{1}{2}\overline{\mu^2}\sqrt{\pi} aL(1-\delta^2)\frac{\omega^2}{c^2}\times$$
$$\times 8\left\{\exp\left(-\frac{l^2}{a^2}\right)\left(1-2\frac{l^2}{a^2}+\frac{1}{2}\frac{l^4}{a^4}\right)\left[\frac{1}{3}\frac{D^2\delta^2}{(1-\delta^2)^2}+\frac{2c^2}{\omega^2 a^2}\frac{1}{(1-\delta^2)^2}\right]\pm\right.$$
$$\left.\pm\frac{1-\delta^2}{2jD}\left[\mathrm{Ei}\left(-\frac{j\frac{l^2}{a^2}}{\frac{D}{1-\delta^2}+j}\right)-\mathrm{Ei}\left(\frac{j\frac{l^2}{a^2}}{\frac{D}{1-\delta^2}-j}\right)\right]\right\}.$$

If the condition $D\delta(1-\delta^2)\ll 1$, which is stronger than the condition $D\delta(1-\delta^2)\gg 1$, is satisfied then by expanding the difference of the E_i functions in accordance with (9) we obtain

$$\overline{\frac{\Delta A_1(s)\,\Delta A_2(s')}{A_0^2}} = \frac{8}{3}\overline{\mu^2}\sqrt{\pi}\left(\frac{L}{a}\right)^3\exp\left(-\frac{l^2}{a^2}\right)\left(1-2\frac{l^2}{a^2}+\frac{1}{2}\frac{l^4}{a^4}\right)\times$$
$$\times\left[1-\frac{3}{8}\frac{1}{\left(\frac{L}{a}\right)^2}\frac{1}{1-\delta^2}\right],$$
$$\overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s')} = \overline{\mu^2}\sqrt{\pi} aL\frac{\omega^2}{c^2}\exp\left(-\frac{l^2}{a^2}\right)(1-\delta^2). \qquad (13)$$

2) $D\delta/(1-\delta^2)\gg 1$. From this condition and from (7) it follows that for large-scale inhomogeneities ($ka\gg 1$) we have

$$b_1(0,\xi) = \frac{1}{1-\delta^2}\left(D\delta-\frac{2c\xi}{\omega a^2}\right)\simeq\frac{D\delta}{1-\delta^2}.$$

If we substitute this value of $b_1(0,\xi)$ into (10), then integration with respect to ξ yields

$$J_1 = \frac{\sqrt{\pi}a^3\omega}{8cj}\frac{1-\delta^2}{\delta}\left\{\mathrm{Ei}\left[-\frac{j\frac{l^2}{a^2}}{\frac{D}{1-\delta^2}+j}\right]-\mathrm{Ei}\left[\frac{j\frac{l^2}{a^2}}{\frac{D\delta}{1-\delta^2}-j}\right]\right\}.$$

Let us substitute the value of J_2 given by (11) and the resultant expression for J_1 into (5). This yields

$$\left.\begin{matrix}\overline{\frac{\Delta A_1(s)\,\Delta A_2(s')}{A_0^2}}\\ \overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s')}\end{matrix}\right\} = \frac{1}{2j}\sqrt{\pi}\,\overline{\mu^2}\left(\frac{a\omega}{2c}\right)^3(1-\delta^2)^2\left\{\frac{1}{\delta}\left[\mathrm{Ei}\left(-\frac{j\frac{l^2}{a^2}}{\frac{D\delta}{1-\delta^2}+j}\right)-\right.\right.$$
$$\left.\left.-\mathrm{Ei}\left(\frac{j\frac{l^2}{a^2}}{\frac{D\delta}{1-\delta^2}-j}\right)\right]\pm\left[\mathrm{Ei}\left(-\frac{j\frac{l^2}{a^2}}{\frac{D}{1-\delta^2}+j}\right)-\mathrm{Ei}\left(\frac{j\frac{l^2}{a^2}}{\frac{D}{1-\delta^2}-j}\right)\right]\right\}. \qquad (14)$$

In this case we can neglect in the arguments of the Ei functions the j terms compared with $D\delta/(1-\delta^2)$ (meaning certainly compared with $D/(1-\delta^2)$).

If we use further the formula Ei (- jx) - Ei(jx) = 2j(π/2) - 2jSi x and put Si x $\simeq$ x for $x \ll 1$, (14) assumes the simple form

$$\left.\frac{\overline{\Delta A_1(s)\,\Delta A_2(s')}}{\overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s')}}\right\} = \frac{1}{4}\sqrt{\pi}\left(\frac{a\omega}{2c}\right)^3\overline{\mu^2}(1-\delta^2)^2\left[\frac{\pi}{2}\left(\frac{1}{\delta}\mp 1\right)-\frac{l^2}{a^2}\frac{(1-\delta^2)(1\mp\delta^2)}{D\delta^2}\right]. \tag{15}$$

Let us note that the conditions $x \ll 1$, that is, $\frac{l^2}{a^2}\frac{1-\delta^2}{D} \ll 1$, for which the expression (15) was derived, signifies that the relative dimensions of the receiving antenna h/a are limited by the requirement $h^2/a^2 \ll D/(1-\delta^2)$.

3. CALCULATION OF THE COEFFICIENTS OF FREQUENCY CORRELATION OF THE AMPLITUDE AND PHASE FLUCTUATIONS AT THE FOCUS OF THE LENS

To obtain the coefficients, defined by formulas (3), of the frequency correlation at the focus of a lens it is necessary to integrate the expressions derived in the foregoing section for the frequency-transverse correlation functions over the area of the lens. This integration will be carried out here for the particular cases $D/(1-\delta^2) \ll 1$ and $D\delta/(1-\delta^2) \gg 1$, when the frequency-transverse correlation functions of the amplitude and phase fluctuations are given by formulas (13) and (15). We shall assume here that the lens is a square with side h (Figure 1).

1) $D/(1-\delta^2) \ll 1$. Under this condition the formulas (13) are valid for $\overline{\Delta A_1(s)\,\Delta A_2(s')}$ and $\overline{\Delta\varphi_1(s)\,\Delta\varphi(s')}$.

Using (13) and (3) we readily obtain the following expressions for the frequency-correlation coefficients

$$R_A^{\Sigma} = 1 - \frac{\frac{3}{8}\frac{1}{\left(\frac{L}{a}\right)^2}\frac{1}{1-\delta^2}}{1-\frac{3}{8}\frac{1}{\left(\frac{L}{a}\right)^2}}, \qquad R_{\varphi}^{\Sigma} = \frac{\frac{3}{8}\frac{1}{\left(\frac{L}{a}\right)^2}}{1-\frac{3}{8}\frac{1}{\left(\frac{L}{a}\right)^2}}, \tag{16}$$

Thus, in this case the correlation coefficients at the focus of the lens are independent of the dimensions of the focusing device, and the phase fluctuations are completely correlated. Formulas (16) coincide with those obtained in [3] for the frequency correlation of the amplitude and phase fluctuations at a point, by putting in the latter $D/(1-\delta^2) \ll 1$.

2. $D\delta/(1-\delta^2) \gg 1$. In this case the values $\overline{\Delta A_1(s)\,\Delta A_2(s')}$ and $\overline{\Delta\varphi_1(s)\,\Delta\varphi_2(s')}$ are determined by formulas (15). We cannot go in these formulas to the limit $\delta = 0$ in order to calculate the denominators of the correlation coefficients (3), since these equations have been obtained under the condition $D\delta/(1-\delta^2) \gg 1$. We shall therefore use formulas (228) and (229) of [5] for the transverse amplitude and phase correlations, namely

$$\left.\frac{\overline{\Delta A_i(s)\,\Delta A_i(s')}}{\overline{\Delta\varphi_i(s)\,\Delta\varphi_i(s')}}\right\} = \frac{1}{2}\overline{\mu^2}\sqrt{\pi}aL\frac{\omega_i^2}{c^2}\left[\exp\left(-\frac{l^2}{a^2}\right)\mp\frac{\pi}{2D_i}\pm\frac{l^2}{a^2}\frac{1}{D_i^2}\right],$$

$$D_i = \frac{4cL}{\omega_i a^2}. \tag{17}$$

Integrating (15) and (17) over the area of the lens and substituting the results in (3) we obtain

$$\left.\begin{matrix} R_A^{\Sigma} \\ R_{\varphi}^{\Sigma} \end{matrix}\right\} = \frac{(1-\delta^2)\left[\frac{\pi}{2D}\left(\frac{1}{\delta}\mp 1\right) - \frac{1}{3}\left(\frac{h}{a}\right)^2 \frac{(1-\delta^2)(1\mp\delta^2)}{D^2\delta^2}\right]}{\left[\frac{Q_0\left(\frac{h}{a}\right)}{h^4} \mp \frac{\pi}{2D}(1-\delta) \pm \frac{1}{3}\left(\frac{h}{a}\right)^2 \frac{(1-\delta)^2}{D^2}\right]^{1/2}} \times$$

$$\times \frac{1}{\left[\frac{Q_0\left(\frac{h}{a}\right)}{h^4} \mp \frac{\pi}{D}(1+\delta) \pm \frac{1}{3}\left(\frac{h}{a}\right)^2 \frac{(1+\delta)^2}{D^2}\right]^{1/2}}, \qquad (18)$$

where the notation $Q_0\left(\frac{h}{a}\right) = \left[\frac{h}{a}\Phi\left(\frac{h}{a}\right) + \frac{\exp\left(-\frac{h^2}{a^2}\right)-1}{\sqrt{\pi}}\right]^2 \pi a^4$, is introduced, and Φ is the probability integral.

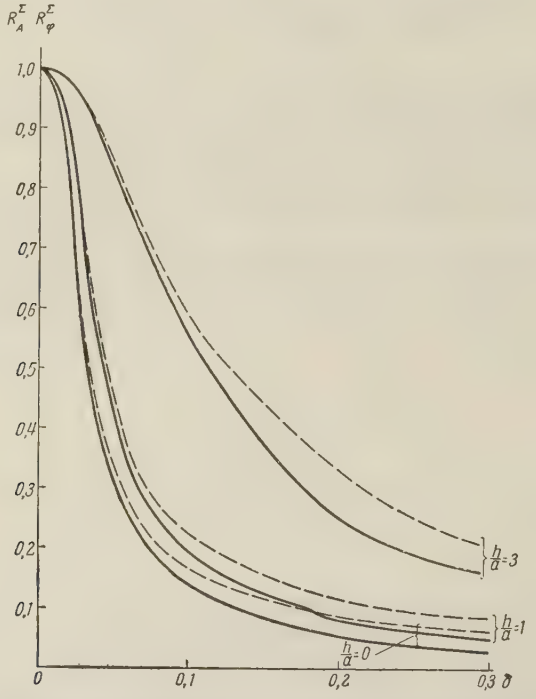


Figure 2. Dependence of the coefficients of frequency correlation of the amplitudes R_A^{Σ} and of the phases R_{φ}^{Σ} on the frequency shift δ for h/a = 0, 1, and 3. Continuous curves — R_A^{Σ}; dashed curves — R_{φ}^{Σ}.

As can be seen from (18), the coefficients of frequency correlation at the focus of the lens depend on the dimension of the receiving apparatus h/a.

In the limiting case, when $h/a \to 0$, formulas (18) go over to the expressions obtained in [3] for the frequency correlation at a point, provided we put in the latter $D\delta/(1-\delta^2) \gg 1$, namely

$$\left.\begin{matrix} \lim\limits_{h\to 0} R_A^{\Sigma} = R_A \\ \lim\limits_{h\to 0} R_{\varphi}^{\Sigma} = R_{\varphi} \end{matrix}\right\} = \frac{(1-\delta^2)\frac{\pi}{2}\left(\frac{1}{\delta}\mp 1\right)}{\sqrt{\left[D\mp\frac{\pi}{2}(1-\delta)\right]\left[D\mp\frac{\pi}{2}(1+\delta)\right]}}.$$

Figure 2 shows plots of the coefficients of frequency correlation of the fluctuations of the amplitude R_A^{Σ} and of the phases R_{φ}^{Σ} at the focus of the lens vs. the frequency shift δ plotted from formulas (18) for lenses with relative dimensions h/a = 0, 1, and 3 and for a wave parameter D = 100. As can be seen from the figure, for the same values of D, h/a, and δ the phase fluctuations are somewhat more correlated than the amplitude fluctuations.

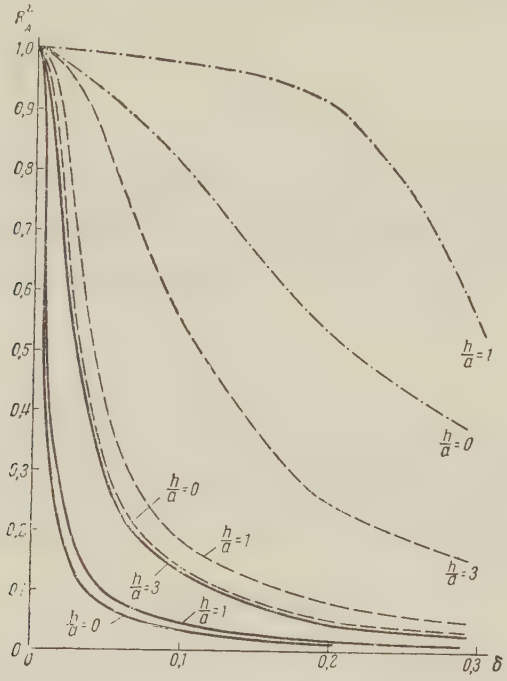


Figure 3. Dependence of the coefficients of frequency correlation of the fluctuations of the amplitude R_A^{Σ} on the frequency shift δ for h/a = 0, 1, and 3; D = 500 (continuous curves); D = 100 (dashes); D = 10 (dash-dots).

Figure 3 shows the dependence of the correlation of the amplitude fluctuations R_A^{Σ} on δ for h/a = 0, 1, 3, D = 500, 100, 10.

It follows from Figures 2 and 3 that as the antenna dimensions are increased, or, what is assumed, as the scattering volume is decreased (since this volume is limited by the directivity pattern of the receiving equipment) the correlation of the amplitudes and phases at the focus increases. At the same time, the frequency-transverse correlation in the antenna edges increases. This agrees with the results of [6], in which it is shown that the decorrelation of the fluctuations of the field in the plane of the antenna aperture is accompanied by a decrease in the scattering volume with increasing antenna dimensions.

The increase of R_A^{Σ} and R_{φ}^{Σ} with increasing h/a signifies also that the absolute fluctuation intensities $\sqrt{\overline{(\Delta A_1^{\Sigma})^2}\,\overline{(\Delta A_2^{\Sigma})^2}}$ and $\sqrt{\overline{(\Delta\varphi_1^{\Sigma})^2}\,\overline{(\Delta\varphi_2^{\Sigma})^2}}$ at the focus of the lens increase more slowly with increasing h/a than the correlation functions $\overline{\Delta A_1^{\Sigma}\Delta A_2^{\Sigma}}$ and $\overline{\Delta\varphi_1^{\Sigma}\Delta\varphi_2^{\Sigma}}$.

In conclusion, the author takes this opportunity to thank S.M. Rytov for a discussion of the work.

REFERENCES

1. L.A. Chernov, Akust. Zh. 1955, 1, 1, 89.

2. L.A. Chernov, Akust. Zh. 1958, 2, 2, 211.
3. M.F. Bakhareva, Radiotekhnika i Elektronika, 1959, 4, 1, 88.
4. V.A. Krasil'nikov, V.I. Tatarskiy, Doklady, Academy of Sciences (New Series) 1953, 38, 3, 435.
5. L.A. Chernov, Propagation of Waves in a Medium with Random Inhomogeneities, Academy of Sciences Press, 1958.
6. M.F. Bakhareva, B.Ye. Kinber, Elektrosvyaz', 1960, 5, 67.

Submitted to the editors, 28 December 1960

EFFECT OF A METALLIC WALL ON MAGNETOSTATIC OSCILLATIONS OF A SMALL GYROTROPIC SPHERE

Hsu Yen-Sheng

The image method is used to examine the effect of a metallic wall on magnetostatic oscillations of a small gyrotropic sphere. It is shown that the effect of the wall is to produce a coupling between the different modes and introduces changes in the resonance conditions. Selection rules are obtained for the interacting modes for the cases of parallel and perpendicular magnetization. The changes in the resonance conditions are calculated for the (1, -1), (2, -2), and (2, -1) modes.

INTRODUCTION

By magnetostatic oscillations we mean the natural oscillations of a small gyrotropic sphere which lturn into the oscillations investigated by Walker [1] when the dimensions of the sphere are decreased. Walker used the magnetostatic approximation to solve the problem. Such oscillations have been studied in [2, 3, 4] under the assumption that the sphere is far from the metal walls of the waveguide or a cavity, and that their effect can be neglected. The proximity of a metallic waveguide wall and its influence on the resonance conditions of uniform precession was accounted for in [5] . However, the method proposed in that reference is complicated, nor were any results obtained for other types of magnetostatic oscillations. In the present paper we use the image method to examine the magnetostatic oscillations of a small gyrotropic sphere, located near a plane metallic wall with infinite conductivity.

1. Calculation of the Influence of a Metallic Wall Using the Image Method

In [2, 3] it has been shown that when magnetostatic oscillations of a small gyrotropic sphere are excited, it is convenient to express the incident and reflected waves in terms of Debye potentials, corresponding to transverse-magnetic and transverse-electric spherical waves. The effect of a metallic wall can be accounted for with the aid of a mirror image of the sphere behind the metallic surface. The fields produced by this sphere are described by additional Debye potentials.

We consider two cases.

1. Case of perpendicular magnetization: the z-axis (that is, the direction of the constant magnetizing field) is perpendicular to the metallic wall (see Figure 1) .

We assume that the center of the sphere coincides with the origin O. The distance from O to the wall is b. Spherical waves, described by Debye potentials u and v, are reflected from the sphere. The influence of the wall can be replaced by spherical waves which are characterized by additional Debye potentials u' and v' with origin O'.

If the fields of the sphere are described by the potentials

$$v = A_{n,m} h_n^{(2)}(k_0 r)\, P_n^m(\cos\theta)\, e^{im\varphi},$$
$$u = B_{n,m} h_n^{(2)}(k_0 r)\, P_n^m(\cos\theta)\, e^{im\varphi}, \tag{1}$$

then the fields of its image are described by the potentials

$$v' = A_{n,m} h_n^{(2)}(k_0 r')\, P_n^m(\cos\theta')\, e^{im\varphi'}$$
$$u' = -B_{n,m} h_n^{(2)}(k_0 r')\, P_n^m(\cos\theta')\, e^{im\varphi'} \tag{2}$$

when n-m is even and

$$v' = -A_{n,m} h_n^{(2)}(k_0 r')\, P_n^m(\cos\theta')\, e^{im\varphi'}$$
$$u' = B_{n,m} h_n^{(2)}(k_0 r')\, P_n^m(\cos\theta')\, e^{im\varphi'} \tag{3}$$

when n-m is odd, where r, θ, φ, and r', θ', φ' are spherical coordinates corresponding to the origins O and O' respectively; $k_0 = \omega/c$ is the wave number in vacuum.

In fact, replacement of the metallic wall by the image of the sphere is valid if the boundary conditions $E_\lambda = 0$ and $H_r = 0$ are satisfied on the wall.

For a TE wave described by a potential v, these boundary conditions are expressed by the following equations on the metallic wall:

$$\frac{ik_0}{r}\frac{\partial}{\partial\theta}(rv) + \frac{ik_0}{r'}\frac{\partial}{\partial\theta'}(r'v') = 0,$$
$$-\frac{ik_0}{r\sin\theta}\frac{\partial}{\partial\varphi}(rv)\cos\theta - \frac{ik_0}{r'\sin\theta'}\frac{\partial}{\partial\varphi'}(r'v')\cos\theta' = 0,$$
$$\left[\frac{\partial^2(rv)}{\partial r^2} + k_0^2 rv\right]\cos\theta + \left[\frac{\partial^2(r'v')}{\partial r'^2} + k_0^2 r'v'\right]\cos\theta' = 0, \tag{4}$$
$$\frac{1}{r}\frac{\partial^2(rv)}{\partial r\,\partial\theta}\sin\theta + \frac{1}{r'}\frac{\partial^2(r'v')}{\partial r'\,\partial\theta'}\sin\theta' = 0.$$

It is obvious that on the wall

$$r = r', \quad \cos\theta = -\cos\theta', \quad \sin\theta = \sin\theta', \quad \varphi = \varphi' \tag{5}$$

and consequently

$$P_n^m(\cos\theta') = P_n^m(\cos\theta)$$
$$\frac{dP_n^m(\cos\theta')}{d\theta'} = -\frac{dP_n^m(\cos\theta)}{d\theta} \qquad \text{when n-m is even}$$
$$P_n^m(\cos\theta') = -P_n^m(\cos\theta) \tag{6}$$
$$\frac{dP_n^m(\cos\theta')}{d\theta'} = \frac{dP_n^m(\cos\theta)}{d\theta} \qquad \text{when n-m is odd}$$

Taking into account the signs in expressions (1) — (3), we can readily verify that condition (4) is satisfied and that out substitution is valid for the Te wave.

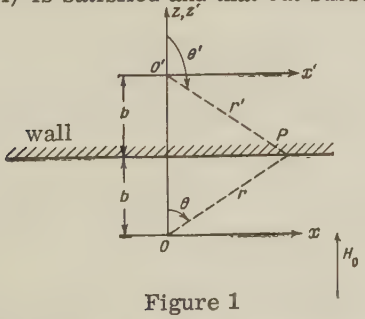


Figure 1

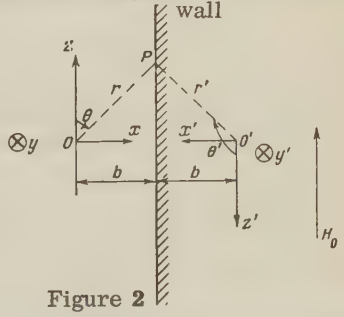


Figure 2

We can show analogously that it is correct to replace the wall by the image of the sphere for the TM wave.

2. Case of parallel magnetization: the z-axis and the magnetizing field are parallel to the metallic wall (see Figure 2).

Unlike the preceding case, the coordinates x' and z', which are associated with the image of the sphere, are directed opposite to the original coordinates x and z. For such coordinate systems it is easy to show that replacement of the metallic wall by the image is also valid,

The rules given above for the choice of the signs and Equations (5) and (5) remain valid in this case, too.

2. COORDINATE-SHIFT FORMULAS FOR THE DEBYE POTENTIALS

We have already indicated in the preceding section that the effect of a metallic wall is equivalent to additional spherical waves, corresponding to Debye potentials u' and v'. Consequently, in order to take into account the influence of the wall, it is necessary to find the formulas for the Debye-potential in the shifted coordinates. Generally speaking this is a very complicated problem. However, the gyrotropic sphere and its distance from the wall are much smaller than the wavelength. In this case the following relation exists between the Debye potential and the magnetostatic and electrostatic potentials ψ'_H and ψ'_E [4]:

$$\psi'_H \simeq \frac{\partial}{\partial r'}(r'v'),$$

$$\psi'_E \simeq \frac{\partial}{\partial r'}(r'u').$$

In the expansion of the Debye potentials u' and v' in powers of the argument $(k_0 r')$ it is sufficient to retain only the first term, and then the problem of finding the coordinate-shift formulas for the Debye potentials u' and v' reduces to the problem of finding the coordinate shifts for the magnetostatic potential ψ'_H and the electrostatic potential ψ'_E, satisfying the Laplace equation.

For the case of perpendicular magnetization, the origin shifts along the z' axis by a distanc -2b, that is, the following coordinate transformation takes place:

$$x = x', \quad y = y', \quad z = z' + 2b.$$

It is known [6] that the following equality holds true:

$$\frac{P_n(\cos\theta')}{r'^{n+1}} = \frac{(-1)^n}{n!}\sum_{s=0}^{\infty}\frac{(n+s)!}{s!}\frac{r^s}{(2b)^{n+s+1}}P_s(\cos\theta). \tag{7}$$

If we apply to both halves of this equation the differential operator $\left[\left(\frac{\partial}{\partial\xi}\right)^m + \left(\frac{\partial}{\partial\eta}\right)^m\right]$, where $\xi = x + iy$; $\eta = x - iy$, we obtain the following equation [7]:

$$\frac{P_n^m(\cos\theta')\,e^{\pm im\varphi'}}{r'^{n+1}} = \frac{(-1)^{n-m}}{(n-m)!}\sum_{s=2m}^{\infty}\frac{(n-m+s)!}{s!}\frac{r^{s-m}P_{s-m}^m(\cos\theta)\,e^{\pm im\varphi}}{(2b)^{n-m+s+1}}. \tag{8}$$

For the case of parallel magnetization, we must first displace the origin along the x' axis by a distance 2b, that is, we must put

$$x'' = x' - 2b, \qquad y'' = y', \qquad z'' = z'.$$

We then must carry out still another transformation:

$$x = -x'', \qquad y = y'', \qquad z = -z''$$

or, in spherical coordinates,

$$r = r'', \quad \theta + \theta'' = 180°, \qquad \varphi + \varphi'' = 180°.$$

The second transformation can be readily carried out, and we shall therefore consider in detail only the first transformation.

Let us consider the expression

$$\frac{1}{r'} = \frac{1}{2b\left[1 + \frac{x''}{b} + \frac{r''^2}{(2b)^2}\right]^{1/2}}. \tag{9}$$

Using the following equation [6]

$$(1 - 2h\mu + h^2)^{-1/2} = P_0(\mu) + P_1(\mu)\,h + P_2(\mu)\,h^2 + \ldots, \tag{10}$$

we can transform (9) into

$$\frac{1}{r'} = \sum_{s=0}^{\infty} \frac{(-r'')^s}{(2b)^{s+1}} P_s(\sin\theta''\cos\varphi''). \tag{11}$$

From the addition formula [6]

$$P_n(\cos\gamma) = P_n(\cos\theta\cos\theta' + \sin\theta\sin\theta'\cos\overline{\varphi - \varphi'}) =$$
$$= P_n(\cos\theta)\,P_n(\cos\theta') +$$
$$+ 2\sum_{m=1}^{n} \frac{(n-m)!}{(n+m)!} P_n^m(\cos\theta)\,P_n^m(\cos\theta')\cos m(\varphi - \varphi')$$

Upon making the substitution

$$\theta = \frac{\pi}{2},\ \varphi = 0$$

it follows that

$$P_n(\sin\theta'\cos\varphi') = P_n(0)\,P_n(\cos\theta') +$$
$$+ 2\sum_{m=1}^{n} \frac{(n-m)!}{(n+m)!} P_n^m(0)\,P_n^m(\cos\theta')\cos m\varphi'.$$

Operating on (11) with the differential operator $\left[\left(\frac{\partial}{\partial\xi}\right)^m \pm \left(\frac{\partial}{\partial\eta}\right)^m\right]\frac{\partial^{n-m}}{\partial z^{n-m}}$, we obtain [6] the required transformation formula

$$(-1)^{n-m}(n-m)!\,\frac{P_n^m(\cos\theta')}{2^m r'^{n+1}} \begin{Bmatrix}\cos m\varphi' \\ -i\sin m\varphi'\end{Bmatrix} =$$
$$= \frac{1}{2}\left[\left(\frac{\partial}{\partial\xi''}\right)^m \pm \left(\frac{\partial}{\partial\eta''}\right)^m\right]\frac{\partial^{n-m}}{\partial z''^{n-m}} \sum_{s=0}^{\infty} \frac{(-r'')^s}{(2b)^{s+1}}\left[P_s(0)\,P_s(\cos\theta'') +\right.$$
$$\left. + 2\sum_{p=1}^{s} \frac{(s-p)!}{(s+p)!} P_s^p(0)\,P_s^p(\cos\theta'')\cos p\varphi''\right]. \tag{12}$$

The right half of (12) can be calculated by using the following formulas (see the appendix):

$$\left(\frac{\partial}{\partial z}\right)^k\left[\left(\frac{\partial}{\partial\xi}\right)^\lambda \pm \left(\frac{\partial}{\partial\eta}\right)^\lambda\right][r^n P_n^m(\cos\theta)\cos m\varphi] =$$
$$= \frac{(-1)^\lambda}{2^\lambda}\frac{(n+m)!}{(n+m-2\lambda-\theta)!} r^{n-\lambda-k} P_{n-\lambda-k}^{m-\lambda}(\cos\theta) \begin{Bmatrix}\cos(m-\lambda)\varphi \\ i\sin(m-\lambda)\varphi\end{Bmatrix} +$$
$$+ \frac{1}{2^\lambda}\frac{(n+m)!}{(n+m-k)!} r^{n-\lambda-k} P_{n-\lambda-k}^{m+\lambda}(\cos\theta) \begin{Bmatrix}\cos(m+\lambda)\varphi \\ -i\sin(m+\lambda)\varphi\end{Bmatrix}$$

when $\lambda \leqslant m$, $k \leqslant n-m$,

$$\left(\frac{\partial}{\partial z}\right)^k\left[\left(\frac{\partial}{\partial\xi}\right)^\lambda \pm \left(\frac{\partial}{\partial\eta}\right)^\lambda\right][r^n P_n^m(\cos\theta)\cos m\varphi] =$$
$$= \frac{(-1)^m}{2^\lambda}\frac{(n+m)!}{(n-m-k)!} r^{n-\lambda-k} P_{n-\lambda-k}^{\lambda-m}(\cos\theta) \begin{Bmatrix}\cos(\lambda-m)\varphi \\ i\sin(m-\lambda)\varphi\end{Bmatrix} +$$
$$+ \frac{1}{2^\lambda}\frac{(n+m)!}{(n+m-k)!} r^{n-\lambda-k} P_{n-\lambda-k}^{\lambda+m}(\cos\theta) \begin{Bmatrix}\cos(m+\lambda)\varphi \\ -i\sin(m+\lambda)\varphi\end{Bmatrix}$$

when $\lambda > m$.

3. EFFECT OF METALLIC WALL ON MAGNETOSTATIC OSCILLATIONS OF A SMALL GYROTROPIC SPHERE

It is seen from the derived formulas that the effect of the wall leads to two phenomena; the shift of the resonant frequency (or of the resonant field value) and a coupling between the oscillation modes, reminiscent of the influence of the effect of wave propagation in a sphere [4]. In the case of perpendicular magnetization, a coupling is produced only between modes having the same index m, while in the case of parallel magnetization the coupling is due to the fact that transformation (8) results only in terms with identical index m, while transformation (12) results only in terms with even or odd index sums n + m. Consequently an interaction should be observed between these oscillation modes, for example (1.-1) and (2, 0) in the case of magnetization parallel to the wall, when the resonant conditions are close to each other (analogous to the interaction between modes (1, -1) and (3, 01) due to the propagation of waves in a sphere [8]).

It should be noted that owing to the coupling between the oscillation modes, it is possible to excite oscillations of one type with the aid of other coupled modes.

A connection exists also between the electrostatic and magnetostatic oscillations, which can be accounted for only in a higher-order approximation when solving the electromagnetic problem. It is small and can be disregarded if $k_0 a \ll 1$ (where a is the radius of the sphere), and consequently $k_0 b \ll 1$. From expressions (8) and (12) we see also that the effect of the wall decreases rapidly with increasing distance from the sphere to the wall. When the ratio b/a is close to unity, that is, the surface of the sphere is almost in contact with the wall, it is necessary to take simultaneous account of the interaction between many higher modes; in this case the problem becomes very complicated. However, it the ratio b/a is greater than 2, the influence of the higher modes is no longer large.

We note that the influence of the walls of a rectangular waveguide or cavity on magnetostatic oscillations of a small gyrotropic sphere can be accounted for by summing the influences of all the walls.

By way of an example let us consider the displacement of the resonant frequency (or resonant value of the magnetizing field) for modes (1, -1) and 2, -2) in the case of magnetization parallel to the wall. Calculation shows that in the first approximation we need account for only the influence of the initial modes, that is, modes (1, -1) and (2, -2) respectively

We consider first homogeneous precession (1, -1). We assume that the wave reflected from the sphere has a Debye potential

$$v_s = C_1 h_1^{(2)}(k_0 r)\, P_1^1(\cos\theta)\, e^{-i\varphi} \simeq \frac{iC_1}{(k_0 r)^2} P_1^1(\cos\theta)\, e^{-i\varphi}$$

when $k_0 r \ll 1$. From the additional Debye potential v'_s we separate the magnetostatic potential of the type (1, -1)

$$\psi'_s = -\frac{iC_1}{k_0^2}\frac{3r}{16b^3} P_1^1(\cos\theta)\, e^{-i\varphi}.$$

Inside the ferrite, on the surface of the sphere, the magnetostatic potential and the normal component of the induction have the form

$$\psi = A_1 P_1^1(\cos\theta)\, e^{-i\varphi}$$
$$B_r = \frac{A_1}{a}(\mu - k) P_1^1(\cos\theta)\, e^{-i\varphi} \qquad \text{for } r = a,$$

where μ and k are the components of the permeability tensor.

From the boundary condition (continuity of the potential and of the normal component B_r of the induction and on the surface of the sphere) we obtain two equations. Equating the determinant of this system to zero, we obtain the resonance condition

$$-\frac{1}{\mu - k - 1} = \frac{1}{3} + \frac{1}{2}\left(\frac{a}{2b}\right)^3 = \Delta_{1,-1}. \tag{14}$$

In [5], a similar problem is solved and the effective permeability μ'_{ef} is solved for homogeneous precession taking into account the losses in the ferrite. If we equate the real part of the denominator of the effective permeability μ_{ef} obtained in that reference, to

zero we obtain a resonance condition which is identical with (14).

When $a/b \to 0$ we get $\Delta_{1,-1} \to 1/3$, which coincides with the well known results obtained by Walker [1]. The term $(a/2b)^3/2$ is a correction to the root $\Delta_{1,-1}$ obtained by Walker, accounting for the influence of the metallic wall on the resonance condition for a sphere of finite radius a.

We can analogously calculate the resonance condition for the (2, -2) mode

$$\Delta_{2,-2} = 0.4 + \frac{7}{4}\left(\frac{a}{2b}\right)^5. \tag{15}$$

By way of another example, let us consider the excitation of a (2, -1) mode by homogeneous precession in perpendicular magnetization. We assume that the wave reflected from the sphere is described by a Debye potential

$$v_s = C_1 h_1^{(2)}(k_0 r)\, P_1^1(\cos\theta)\, e^{-i\varphi} + C_2 h_2^{(2)}(k_0 r)\, P_2^1(\cos\theta)\, e^{-i\varphi} \cong$$
$$\cong \frac{iC_1}{(k_0 r)^2} P_1^1(\cos\theta)\, e^{-i\varphi} + \frac{i3C_2 P_2^1(\cos\theta)}{(k_0 r)^3} e^{-i\varphi}$$

for $k_0 r \ll 1$.

From the additional Debye potential v'_s we separate the magnetostatic potential

$$\psi'_s = \frac{-iC_1}{k_0^2}\left[\frac{rP_1^1(\cos\theta)\, e^{-i\varphi}}{(2b)^3} + \frac{r^2 P_2^1(\cos\theta)\, e^{-i\varphi}}{(2b)^4}\right] +$$
$$+ \frac{6iC_2}{k_0^3}\left[\frac{3rP_1^1(\cos\theta)\, e^{-i\varphi}}{(2b)^4} + \frac{4r^2 P_2^1(\cos\theta)\, e^{-i\varphi}}{(2b)^5}\right].$$

Inside the ferrite on the surface of the sphere, the potential has the form

$$\psi = A_1 P_1^1(\cos\theta)\, e^{-i\varphi} + A_2 P_2^1(\cos\theta)\, e^{-i\varphi}.$$

The boundary conditions lead to a system of four equations. Setting its determinant equal to zero, we obtain the resonance conditions

$$\mu - k - 1 \simeq -3 + 3\left(\frac{a}{2b}\right)^3,$$
$$\mu - k - 1 \simeq -5 - 20\left(\frac{a}{2b}\right)^5. \tag{16}$$

The first condition corresponds to resonance of mode (1, -1), and the second to the resonance of the (2, -1) mode. As $a/b \to 0$ these expressions yield $\Delta_{1,-1} = 1/3$ and $\Delta_{2,-1} = 1/5$, which coincides with [1]. The ratio A_1, A_2 determines the relative amplitude of the (1, -1) and 2, -1) modes in the sphere.

CONCLUSION

The effect of a metallic wall causes a displacement of the resonant frequency (or the resonant value of the magnetizing field) and results in a coupling between various oscillation modes. This effect decreases rapidly with increasing distance from the sphere to the wall and can be neglected in practice if $b/a > 3$. Calculation shows that a coupling exists between modes with like indices m in the case of magnetization perpendicular to the wall, and between modes with even or odd index sums n+m in the case of magnetization parallel to the wall. The formulas given above enable us to calculate corrections for the resonance conditions and to solve the problem of exciting one mode by another mode coupled to it.

I am grateful to A.A. Pistol'kors for guidance and to A.V. Vashkovskiy for many valuable suggestions.

APPENDIX

Derivation of Formulas (13)

Formulas (13) with the upper signs and cosinusoidal dependence on φ are found in

Hobson's book. It is necessary to derive these formulas for the lower signs and for a sinusoidal dependence.

It is easy to show that

$$\frac{(\xi^m + \eta^m)}{2} = r^m \sin^m \theta \cos m\varphi,$$

$$\frac{(\xi^m - \eta^m)}{2} = r^m \sin^m \theta \sin m\varphi,$$

from which it follows that

$$r^n P_n^m(\cos\theta)\sin m\varphi = (-1)^m \frac{(n+m)!}{2^m (n-m)!\, m!}\left[Z^{n-m} - \frac{(n-m)(n-m-1)}{2(2m+2)} Z^{n-m-2}\xi\eta + \right.$$

$$\left. + \frac{(n-m)\ldots(n-m-3)}{2\cdot 4\cdot (2m+2)(2m+4)} Z^{n-m-4}(\xi\eta)^2 + \ldots \right] \frac{(\xi^m - \eta^m)}{2i}.$$

With the aid of a transformation method similar to the developed in Section 88 of [6] we can obtain formulas (13).

REFERENCES

1. L.R. Walker, Phys. Rev. 1957, 105, 2, 390.
2. A.A. Pistol'kors, Hsu Yen-Sheng, Radiotekhnika i Elektronika, 1960, 5, 1, 3.
3. Hsu Yen-Sheng, Radiotekhnika i Elektronika, 1960, 5, 1, 15.
4. Hsu Yen-Sheng, Radiotekhnika i Elektronika, 1960, 5, 12, 1951.
5. W. Hauser, L. Brown, J. Appl. Phys, 1959, 30, 9, 1460.
6. E.W. Hobson, The Theory of Spherical and Ellipsoidal Functions, Cambridge University Press, 1931.
7. L. Robin, Fonctions spheriques de Legendre et fonctions spheroidals, 1, 1957.
8. P. Fletcher, I.H. Solt, J. Appl. Phys., 1959, Suppl. to 30, 4, 181 s.

Submitted to the editors 19 January 1961

DIFFRACTION OF ELECTROMAGNETIC WAVES ON A CONCAVE SPHERICAL SURFACE

B. Ye. Kinber

The axially symmetrical problem of diffraction of a toroidal wave, excited by an annular current filament, on the internal surface of a sphere is considered. The condition of periodicity in the angle θ is replaced by the radiation condition. An investigation of the short-wave asymptotic approximation of the solution shows that the resultant field can be represented as a sum of rays of multiply reflected waves, the caustic fields, and the field of the "whispering gallery" wave. The difference between diffraction on a sphere and the diffraction on a cylinder [1] is determined by the focusing factor, which has a purely geometrical meaning away from the symmetry axis and a purely diffraction meaning near the axis. The problem investigated makes it possible to describe the diffraction of an edge wave on the concave side of a parabolic antenna.

INTRODUCTION

When a wave strikes a reflector of finite dimensions, a cylindrical or toroidal wave is propagated from its edge; this has been shown by an analysis of the solution of the diffraction

problem on a half-plane and on a disc. Owing to the concave surface of the reflector, this wave is reflected several times from the surface of the reflector.

In [1] we investigated certain laws of diffraction of such a wave on a concave cylindrical surface. We consider below the laws of diffraction near a concave surface of double curvature, the simplest of which, namely a spherical surface, was chosen. To simplify the problem, we consider the axially symmetrical field excited by a current ring.

In a real reflector the waves reflected from its concave surface move away from the reflector and the field satisfies the radiation condition. In the problem we are considering here, to simplify the solution, we chose a closed spherical surface, and the radiation condition is introduced beforehand, that is, the solution is sought outside the ring in the form of a sum of normal waves, "running away" in an angle θ from the ring. This form of solution was used earlier [2, 3] in an analysis of the propagation of super-long waves between the earth's surface and the ionosphere. Thus, unlike the ordinary definition in a spherical coordinate system, θ varies in the interval $(-\infty, \infty)$. Values of θ differing by $2\pi n$ correspond to a different number of "circuits" of the wave, that is, to different phase factors. The range of variation of θ which is of interest to us, taking the symmetry of the problem into account, lies in the interval $(0, \pi)$.

Naturally, if the problem has axial symmetry the location of the rays in the meridional plane of the sphere coincides with the location of the rays in a cylinder. Therfore some of the derivations, which are the same as in the analogous calculations for the cylinder, are not given here.

1. EIGENFUNCTIONS OF THE PROBLEM

The solution of the axially symmetrical problem can be represented in the form of a complete system of eigenfunctions satisfying the boundary conditions. All the components can be expressed in terms of an Abraham potential which has only a component of the field, $\vec{A} = \vec{\varphi}_0 A(\theta, R)$ for both electric fields

$$\begin{aligned} H_R &= -\frac{1}{\omega\mu\varepsilon R\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta A), \quad E_R = 0, \\ H_\theta &= \frac{1}{\omega\mu\varepsilon R}\frac{\partial}{\partial R}(RA), \quad E_\theta = 0, \\ H_\varphi &= 0, \quad E_\varphi = i\omega A, \end{aligned} \tag{1}$$

and for magnetic fields

$$\begin{aligned} H_R &= 0, \quad E_R = -\frac{1}{\omega\mu\varepsilon R\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta A), \\ H_\theta &= 0, \quad E_\theta = \frac{1}{\omega\mu\varepsilon R\sin\theta}\frac{\partial}{\partial R}(RA), \\ H_\varphi &= -i\omega A, \quad E_\varphi = 0. \end{aligned} \tag{2}$$

The potential A satisfies the equation

$$\frac{1}{R}\frac{\partial^2}{\partial R^2}(RA) + \frac{1}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta A)\right) + K^2 A = 0, \tag{3}$$

where $k^2 = \omega^2\mu\epsilon$.

The boundary condition for the potential has the form

$$A = 0, \tag{4}$$

for electric waves and

$$\frac{\partial}{\partial R}(RA) = 0. \tag{5}$$

for magnetic waves.

Equation (3) is solved by separating the variables

$$A = U(R)\,\Phi(\theta). \tag{6}$$

The solution for U with no singularity at the center of the sphere has the form

$$U = \frac{1}{\sqrt{kR}} J_{\nu+\frac{1}{2}}(kR), \quad \operatorname{Re}\nu \geqslant 0, \quad \operatorname{Im}\nu \geqslant 0, \tag{7}$$

and the eigenvalues ν are determined from (4) or (5) as the roots of the equations

$$J_{\nu+\frac{1}{2}}(ka) = 0, \tag{4a}$$

$$\frac{\partial}{\partial(kR)}\left[\sqrt{kR}J_{\nu+\frac{1}{2}}(kR)\right]\Big|_{R=a} = 0, \tag{5a}$$

where a is the radius of the sphere.

We choose the solution for Φ in the form $P^1_\nu(\cos\theta)$ or in the form

$$L_\nu = Q^1_\nu(\cos\theta) + i\frac{\pi}{2} P^1_\nu(\cos\theta).$$

The real roots of Equations (4a) or (5a) correspond to propagating waves, while the imaginary ones correspond to damped waves which describe the singularity of the source.

2. GREEN'S FUNCTION

Let us define the Green's function $\Gamma = \Gamma(R, \theta, R', \theta')$ of the problem, that is, a function satisfying the following conditions.

1. Γ satisfies the wave equation (3) everywhere except in the ring (R', θ') ($kR' \sin\theta \gg 1$), where the current filament is situated.

2. Near the ring (R', θ'), the function Γ has a singularity

$$\Gamma = \pi i H^{(1)}_0(k\rho) + \Gamma' \cong -2\ln k\rho, \tag{8}$$

where ρ is the distance from the point of observation to the nearest point of the ring and Γ' is the function with no singularity.

3. As $|\theta - \theta'| \to \infty$, the function Γ satisfies a radiation condition of the waveguide type, that is, it is a wave traveling from the source.

4. On the surface R = a of the sphere, the function Γ satisfies the boundary condition (4a) or (5a), and has no singularities at the center of the sphere.

We use further in lieu of the excitation condition (8) its equivalent

$$\frac{\partial\Gamma}{\partial\theta}\Big|_{\theta=\theta'+0} - \frac{\partial\Gamma}{\partial\theta}\Big|_{\theta=\theta'-0} = -4\pi R\delta(R - R') \tag{9}$$

on the cone $\theta = \theta'$.

We shall seek Γ in the form

$$\Gamma = \frac{1}{\sqrt{kR}}\sum_s A_s J_{\nu_s+\frac{1}{2}}(kR)\, L_{\nu_s}(\cos\theta)\, P^1_{\nu_s}(\cos\theta'), \quad \theta \geqslant \theta',$$

$$\Gamma = \frac{1}{\sqrt{kR}}\sum_s A_s J_{\nu_s+\frac{1}{2}}(kR)\, P^1_{\nu_s}(\cos\theta)\, L_{\nu_s}(\cos\theta'), \quad \theta \leqslant \theta'. \tag{10}$$

The summation is over all the eigenvalues ν_s.

Substituting (10) into (9) and recognizing that

$$\frac{d}{d\theta} L_\nu P^1_\nu - \frac{d}{d\theta} P^1_\nu L_\nu = -\frac{\nu(\nu+1)}{\sin\theta}, \tag{11}$$

Multiplying both halves of the equation by $\frac{1}{\sqrt{kR}} J_{\nu_s+\frac{1}{2}}(kR)$, integrating with respect to R over the limits (0, a), using the relation

$$\int \frac{J_n J_m}{x}\, dx = \frac{x}{n^2 - m^2}\,[J_n J'_m - J_m J'_n] \qquad (x = kR)$$

and the boundary conditions, we obtain

$$A_s = 8\pi \frac{\sqrt{kR'}}{ka} \sin\theta' \frac{\nu_s + \frac{1}{2}}{\nu_s(\nu_s+1)} J_{\nu_s+\frac{1}{2}}(kR') \frac{1}{D_s}, \tag{12}$$

where

$$D_s = \frac{1}{\sqrt{kR}} \frac{d}{d\nu} J_{\nu+\frac{1}{2}}(kR) \frac{d}{d(kR)} \left[\sqrt{kR}\, J_{\nu+\frac{1}{2}}(kR)\right]\Big|_{\substack{R=a \\ \nu=\nu_s}} \tag{13a}$$

for boundary condition (4a) and

$$D_s = -\frac{1}{\sqrt{kR}} J_{\nu+\frac{1}{2}}(kR) \frac{d^2}{d\nu\, d(kR)} \left[\sqrt{kR}\, J_{\nu+\frac{1}{2}}(kR)\right]\Big|_{\substack{R=a \\ \nu=\nu_s}} \tag{13b}$$

for boundary condition (5a).

Substituting (12) in (10) we obtain

$$\Gamma = \frac{8\pi}{ka} \sqrt{\frac{R'}{R}} \sin\theta' \sum_s \frac{\nu_s + \frac{1}{2}}{\nu_s(\nu_s+1) D_s} J_{\nu_s+\frac{1}{2}}(kR) J_{\nu_s+\frac{1}{2}}(kR') \begin{Bmatrix} L_{\nu_s} & P^1_{\nu_s} \\ P^1_{\nu_s} & L_{\nu_s} \end{Bmatrix}, \tag{14}$$

where for $\theta \geqslant \theta'$ we take the upper line in the braces, and for $\theta \leqslant \theta'$ we take the lower line.

We consider further the boundary condition (5a), and place the current ring on a wall. In this case expression (14) can be represented in the form of an integral along a contour C_0 which includes all the eigenvalues ν_s on the complex plane ν, but does not include the origin

$$\Gamma = \frac{L}{ka} \sqrt{\frac{a}{R}} \sin\theta' \int_{C_0} \frac{J_{\nu+\frac{1}{2}}(kR)}{J'_{\nu+\frac{1}{2}}(ka) + \frac{1}{2ka} J_{\nu+\frac{1}{2}}(ka)} \frac{\nu+\frac{1}{2}}{\nu(\nu+1)} \times$$

$$\times \begin{Bmatrix} L_\nu(\cos\theta) & P^1_\nu(\cos\theta') \\ P^1_\nu(\cos\theta) & L_\nu(\cos\theta') \end{Bmatrix} d\nu. \tag{15}$$

3. ASYMPTOTIC EXPRESSIONS FOR THE SOLUTION WHEN $ka \gg 1$

We consider the asymptotic expression for solution (15) for a sphere of large radius, $ka \gg 1$. In this case we can discard the term $\frac{1}{2ka} J_{\nu+\frac{1}{2}}(ka)$ in the denominator.

In accordance with the asymptotic representation of the associated Legendre functions, we shall distinguish between two cases for the values of ν at the stationary-phase points (see [1[):

a) $\theta\nu \gg 1$ — region away from the symmetry axis,

b) $\theta\nu \leqslant 1$ — (or $(\pi - \theta)\nu \leqslant 1$) — region of the focal point near the symmetry axis.

Let us examine both cases.

a) $\theta\nu \gg 1$.

In this case

$$L_\nu = \sqrt{\frac{\pi\nu}{2\sin\theta}}\exp\left\{i\left[\left(\nu+\frac{1}{2}\right)\theta+\frac{3\pi}{4}\right]\right\},$$
$$P_\nu^1 = \sqrt{\frac{2\nu}{\pi\sin\theta}}\sin\left[\left(\nu+\frac{1}{2}\right)\theta+\frac{3\pi}{4}\right] \tag{16}$$

and (15) can be written in the form

$$\Gamma = -\frac{2}{ka}\sqrt{\frac{a\sin\theta'}{R\sin\theta}}\int_{\nu+\frac{1}{2}(ka)}^{J_{\nu+\frac{1}{2}}(kR)}\left\{e^{i\left[\left(\nu+\frac{1}{2}\right)(\theta+\theta')\frac{\pi}{2}\right]} - e^{\mp i\left(\nu+\frac{1}{2}\right)(\theta-\theta')}\right\}d\nu, \tag{17}$$

The plus sign in the exponent of the second term corresponds to the case $\theta > \theta'$, and the minus sign to the case $\theta < \theta'$.

It was assumed in the derivation that $\nu \gg 1$, that is, $\left(\nu+\frac{1}{2}\right)\Big/(\nu+1) = 1$. It is easy to see that the factor $\sqrt{a\sin\theta'/R\sin\theta}$ corresponds to the purely geometrical amplification factor of the amplitude, due to the reduction in the cross section of the ray tubes between the meridional planes.

In all other respects formula (17) corresponds to formula (19) of [1], if we take additional account of the fact that the term with phase $\left(\nu+\frac{1}{2}\right)(\theta+\theta')-\frac{\pi}{2}$ corresponds to the radiating point on the far side of the current ring, while the term with the phase $\mp\left(\nu+\frac{1}{2}\right)(\theta-\theta')-$ corresponds to the radiation from the near point of the current ring. The additional phase $-\pi/2$ in the first term is connected with the passage of the field through the axial caustic.

Thus, the difference between the diffraction of a current filament on a concave cylindrical surface and that of a current ring on a spherical surface away from the axis lies only in the additional geometrical focusing, connected with the axial symmetry of the problem and with the presence of two "sources" corresponding to the far and near points of the ring. The total field can be represented as a sum of rays passing through or past the point of observation, and the fields of surface waves radiated by two sources, namely the near and far points of observation.

Inside the current ring, the rays and waves move opposite each other and form a standing wave; outside the ring, they move approximately in one direction.

The number of waves, their amplitude, and the placement of the caustics correspond to the analysis made for the cylindrical surface.

b) $\theta\nu \leqslant 1$. The case $\theta\nu \leqslant 1$ corresponds to the case of interest to us, namely that of diffraction of a fringe wave from the edge of a reflector in the center of the concave surface of the reflector. Since we have assumed before hand that the diameter of the current ring is much greater than the wavelength, the asymptotic expression for $L_\nu(\cos\theta)$ is taken in (15) in accordance with (16) and the following expression holds true for $P_\nu^1(\cos\theta)$ [4]:

$$P_\nu^1(\cos\theta) \simeq -\left(\nu+\frac{1}{2}\right)\sqrt{\frac{\theta}{\sin\theta}}J_1\left[\left(\nu+\frac{1}{2}\right)\theta\right]. \tag{18}$$

Consequently,

$$\Gamma = i\frac{4\sqrt{\frac{\pi}{2}}}{ka}\sqrt{\frac{a\sin\theta'}{R}\frac{\theta}{\sin\theta}}\int\sqrt{\nu}J_1\left[\left(\nu+\frac{1}{2}\right)\theta\right]\times$$
$$\times\frac{J_{\nu+\frac{1}{2}}(kR)}{J_{\nu+\frac{1}{2}}(ka)}e^{i\left[\left(\nu+\frac{1}{2}\right)\theta'+\frac{3\pi}{4}\right]}d\nu. \tag{19}$$

Inasmuch as the function (18) varies slowly compared with the function $J_{\nu+\frac{1}{2}}(kR) \Big/ J'_{\nu+\frac{1}{2}}(ka)$ as ν is varied, we can regard this function as slowly varying in the integration by the stationary-phase method.

In all other respects the structure of the integrand in (19) corresponds to the integrand of the first term of (17) with $\theta = 0$, and consequently the integral (19) yields the ordinary amplitudes of the ray and caustic field, multiplied by a factor $\sqrt{\nu} J_1\left[\left(\nu + \frac{1}{2}\right)\theta'\right]$ which can be regarded as the diffraction focusing factor near the focal point.

As noted earlier, [1], at the stationary) phase points ν should be smaller than ka in order to be able to use the stationary-phase method for the integration. It is therefore necessary to transform into an integral not the entire series (14), but to leave in the old form the normal waves with maximum ν_s, which are interpreted as a whispering-gallery surface wave. It is easy to see that the analysis given above for the asymptotics of the angle functions is applicable to the series as it is to the integral. Consequently, along with axial focusing of the rays and caustics, the surface wave is also focused.

CONCLUSION

An analysis of the asymptotic behavior of the diffraction on a double-curvature surface has shown that the difference between the diffraction on such a surface and diffraction on a cylindrical surface lies in the geometrical-focusing factor away from the symmetry axis and in the diffraction-focusing factor near the symmetry axis, with respect to all the elements making up the total field — rays, caustics, and surface waves.

REFERENCES

1. B.Ye. Kinber, On the Diffraction of Electromagnetic Waves on a Concave Surface of a Circular Cylinder, Radiotekhnika i Elektronika, 1961, 6, 8, 1273.
2. P.E. Krasnushkin, Method of Normal Waves as Applied to Long-Distance Radio Communication, Press of the Moscow State University, 1947.
3. W. Schuman, Uber die Oberfelder bei der Ausbreitung langer electrischer Wellen, Z. angew. Phys., 1954, 6, p. 35, 225, 267, 346.
4. L.A. Vaynshteyn, Electromagnetic Waves, Soviet Radio Press, 1955.

Submitted to the editor 1 December 1960

DIFFRACTION OF SURFACE ELECTROMAGNETIC WAVES ON AN IMPEDANCE STEP OF A CIRCULAR CYLINDER

V.V. Kolpakov

The problem of the diffraction of axially symmetrical surface waves on an impedance step of a circular cylinder is considered. Expressions for the reflection and transmission power coefficients are obtained, as well as radiation diagrams.

INTRODUCTION

Trenev [1], Weill [2], and Kay [3] have considered the diffraction of surface electromagnetic waves on an impedance step, specified on an infinite plane. An analogous problem

was discussed by Talanov [4] as the limiting case of the diffraction of electromagnetic waves by an impedance step of one of the walls of a flat waveguide.

In the present paper we consider the diffraction of electromagnetic surface waves on an impedance step for a more complicated system — an unbounded impedance cylinder. Since cylindrical systems are used along with plane surface-waves systems, a solution of such a problem is of both theoretical and practical interest.

1. FORMULATION OF THE PROBLEM

Let us examine a surface axially symmetrical TM wave. The solution of the problem reduces to the integration of the differential equation

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial H}{\partial\rho}\right)+\frac{\partial^2 H}{\partial z^2}+\left(k^2-\frac{1}{\rho^2}\right)H=0$$

for a single magnetic-field component $H_\varphi = H(\rho, z)$, using the boundary conditions on the surface of the impedance cylinder

$$E_z = ZH_\varphi \tag{1}$$

and the conditions at infinity. In the plane z = 0 an impedance jump takes place

$$Z=\begin{cases}Z_1=-iQ_1 & \text{for } z<0,\\ Z_2=-iQ_2 & \text{for } z>0.\end{cases} \tag{2}$$

The components of the electric field are found from the equations

$$E_\rho=-\frac{i}{k}\frac{\partial H}{\partial z},\qquad E_z=\frac{i}{k}\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho H), \tag{3}$$

where $k = \omega/c = 2\pi\lambda$. The time dependence is taken in the form exp ($-i\omega t$).

Taking relations (2) and (3) into account, the boundary conditions (1) assume the form

$$\frac{\partial H}{\partial\rho}+\frac{akQ+1}{a}H=0 \qquad \text{for } \rho=a. \tag{4}$$

To the left ($z < 0$) there is incident on the impedance step a surface wave

$$H(\rho, z) = H_1^{(1)}(\rho, v_1)\, e^{ih_1 z}, \tag{5}$$

where the propagation constant h_1 is defined by the dispersion equation

$$v_1 H_0^{(1)}(av_1) + kQ_1 H_1^{(1)}(av_1) = 0, \qquad v_1 = \sqrt{k^2 - h_1^2}, \tag{6}$$

which can be obtained from (4) and (5).

2. SOLUTION OF THE INTEGRAL EQUATION

With the aid of Green's theorem, let us write down the solution of the problem in integral form. We stipulate here that the Green's function satisfy the boundary conditions (4) for $Z = Z_1$ over the entire length of the cylinder. As a result we obtain

$$H(\rho, z) = H_1^{(1)}(\rho, v_1)\, e^{ih_1 z} + 2\pi ak\,(Q_2 - Q_1)\int_0^\infty G(\rho, a, z, z')\, H(a, z')\, dz'. \tag{7}$$

The Fourier amplitude $g(\rho, a, w)$ of the Green's function $G(\rho, a, z, z')$ is found in the form

$$g(\rho, a, w) = -\frac{1}{2\pi a}\frac{H_1^{(1)}(\rho v)}{vH_0^{(1)}(av) + kQ_1H_1^{(1)}(av)}, \quad \rho \geqslant a, \tag{8}$$

where $v = \sqrt{k^2 - w^2}$.

Thus, in order to calculate the field components from formulas (7) and (3) it is necessary first to calculate the tangential component of the magnetic field H(a, z) for $z > 0$. Putting in (7) $\rho = a$ and $z > 0$, we obtain the integral equation for this component

$$H(a, z) = H_1^{(1)}(av_1)\, e^{ih_1 z} + 2\pi ak\,(Q_2 - Q_1)\int_0^\infty G(a, a, z, z')\, H(a, z')\, dz'. \tag{9}$$

The resultant inhomogeneous integral equation (9) with symmetrical kernel is called the Wiener-Hopf equation. Methods of solving this equation have by now been sufficiently well developed [5]. The equation is solved on the complex plane. For this purpose we assume for the time being that the propagation constant k has a vanishingly small positive imaginary part. Leaving out the details of the calculation, let us write down the result of solving Equation (9). The Fourier amplitude h(a, w) of the sought function H(a, z) has a form

$$h(a, w) = \int_0^\infty e^{iwz} H(a, z)\, dz = iH_1^{(1)}(av_1)\,\Psi_1(h_1)\frac{\Psi_1(w)}{w + h_1}, \tag{10}$$

where the functions

$$\Psi_1(w) = \sqrt{\frac{[vH_0^{(1)}(av) + kQ_1H_1^{(1)}(av)]\,(w + h_1)\,(w - h_2)}{[vH_0^{(1)}(av) + kQ_2H_1^{(1)}(av)]\,(w - h_1)\,(w + h_2)}}\exp\left[\frac{M(w)}{2}\right],$$

$$\Psi_2(w) = \Psi_1(-w) = \sqrt{\frac{[vH_0^{(1)}(av) + kQ_1H_1^{(1)}(av)]\,(w - h_1)\,(w + h_2)}{[vH_0^{(1)}(av) + kQ_2H_1^{(1)}(av)]\,(w + h_1)\,(w - h_2)}}\exp\left[-\frac{M(w)}{2}\right] \tag{11}$$

are the result of the expansion

$$\frac{vH_0^{(1)}(av) + kQ_1H_1^{(1)}(av)}{vH_0^{(1)}(av) + kQ_2H_1^{(1)}(av)} = \Psi_1(w)\,\Psi_2(w). \tag{12}$$

The numbers h_1 and h_2 are the propagation constants of the surface waves for $z < 0$ and $z > 0$, respectively; h_1 is a solution of the dispersion equation (6), and h_2 is a solution of the equation

$$v_2H_0^{(1)}(av_2) + kQ_2H_1^{(1)}(av_2) = 0, \quad v_2 = \sqrt{k^2 - h_2^2}. \tag{13}$$

The functions $\Psi_1(w)$ and $\Psi_2(w)$ have the following properties. The function $\Psi_1(w)$ is holomorphic, has no zeros in the half-plane Im $W > -c$, and becomes equal to unity in this half-plane as $|w| \to \infty$; the function $\Psi_1(w)$ has the same property in the plane Im $w < c$, where $c \leqslant \text{Im } h_1$ and $c \leqslant \text{Im } h_2$. The function

$$M(u) = X(u) + iY(u)$$

has no singularities, and

$$X(u) = \frac{2u}{\pi}\int_0^k \frac{\Omega(w)\,dw}{w^2 - u^2}$$

$$Y(u) = -\frac{1}{i\pi}\int_{-i\infty}^{+i\infty}\frac{\Omega(w)\,dw}{w - u} \qquad \text{when } |u| > k; \tag{14}$$

$$X(u) = \frac{2u}{\pi} \int_0^k \frac{v\,\Omega(v) - v_0\Omega(v_0)}{(v_0^2 - v^2)\sqrt{k^2 - v^2}}\,dv \qquad \text{when } |u| < k; \tag{15}$$

$$Y(u) = -\Omega(v_0) - \frac{1}{i\pi} \int_{-i\infty}^{+i\infty} \frac{\Omega(w)\,dw}{w - u}$$

$$v_0 = \sqrt{k^2 - u^2};$$

$$\Omega(w) = \operatorname{arctg} \frac{vN_0(av) + kQ_1N_1(av)}{vJ_0(av) + kQ_1N(av)} - \operatorname{arctg} \frac{vN_0(av) + kQ_2N_1(av)}{vJ_0(av) + kQ_2J_1(av)}. \tag{16}$$

Taking the Fourier transform of (7), and using at the same time relations (8) and (9), we get

$$H(\rho, z) = H_1^{(1)}(\rho v_1)\, e^{ih_1 z} +$$

$$+ \frac{k(Q_2 - Q_1)}{2\pi i} H_1^{(1)}(av_1)\,\Psi_1(h_1) \int_{-\infty}^{+\infty} e^{-iwz} \frac{\Psi_1(w)\, H_1^{(1)}(\rho v)\,dw}{(w + h_1)\,[vH_0^{(1)}(av) + kQ_2H_1^{(1)}(av)]}. \tag{17}$$

With the aid of (12) we reduce (17) to the form

$$H(\rho, z) = H_1^{(1)}(\rho v_1)\, e^{ihz} +$$

$$+ \frac{k(Q_2 - Q_1)}{2\pi i} H_1^{(1)}(av_1)\,\Psi_1(h_1) \int_{-\infty}^{+\infty} e^{-iwz} \frac{H_1^{(1)}(\rho v)\,dw}{(w + h_1)\,\Psi_2(w)\,[vH_0^{(1)}(av) + kQ_2H_1^{(1)}(av)]}. \tag{18}$$

3. CALCULATION OF THE FIELDS

In order to calculate the field of the surface wave for $z < 0$, we deform the integration path in (18) in the upper half-plane Im $w > 0$. Calculating the residue at the pole $w = h_1$, using at the same time formulas (11), we obtain

$$\underset{i}{H}(\rho, z) = H_1^{(1)}(\rho v_1)\, e^{ih_1 z} + \frac{h_1 - h_2}{h_1 + h_2} \exp M(h_1)\, H_1^{(1)}(\rho v_1)\, e^{-ih_1 z}, \quad z < 0. \tag{19a}$$

Here the first term is the surface wave incoming on the step, and the second term is the reflected wave.

Deforming for $z > 0$ in (17) the integration path in the lower half-plane Im $w < 0$ and calculating the residues at the poles $w = -h_1$ and $w = -h_2$, we obtain the field of the transmitted surface wave

$$H(\rho, z) = \frac{2h_1}{h_1 + h_2} \sqrt{\frac{[H_1^{(1)}(av_1)]^2 - H_0^{(1)}(av_1)\, H_2^{(1)}(av_1)}{[H_1^{(1)}(av_2)]^2 - H_0^{(1)}(av_2)\, H_2^{(1)}(av)}} \times$$

$$\times \exp\left[\frac{M(h_1)}{2} - \frac{M(h_2)}{2}\right] H_1^{(1)}(\rho v_2)\, e^{ih_2 z}, \qquad z > 0. \tag{19b}$$

The reflection and transmission power coefficients R and T are determined by

$$R = \left(\frac{\beta_1 - \beta_2}{\beta_1 + \beta_2}\right)^2 \exp 2X(\beta_1), \tag{20}$$

$$T = \frac{4\beta_1\beta_2}{(\beta_1 + \beta_2)^2} \exp\,[X(\beta_1) - X(\beta_2)],$$

where $\beta_1 = h_1/k$ and $\beta_2 = h_2/k$ are the retardation coefficients to the right and to the left of the impedance step.

Formulas (20) have been obtained by calculating the power flux of the incoming, reflecting, and transmitted waves in the plane z = const to the right and to the left of the impedance step. Inasmuch as $\beta_1 > 1$ and $\beta_2 > 1$, we calculate $X(\beta_1)$ and $X(\beta_2)$ in (20) by means of formula (14).

It is interesting to trace the dependence of the reflection and transmission coefficients on a/λ at large values of ak. For these values of ak we can represent the cylindrical functions in (7) (13) and (16 by semiconvergent series. Expanding the function $\Omega(w)$ in powers of $1/ak$ and discarding terms beyond the second power of $1/ak$, we obtain after rather cumbersome but simple calculations

$$\Omega(w) = \operatorname{arctg}\frac{k\sqrt{\beta_2^2-1}}{v} - \operatorname{arctg}\frac{k\sqrt{\beta_1^2-1}}{v} + \frac{3}{8a^2kv}\frac{\sqrt{\beta_1^2-1}-\sqrt{\beta_2^2-1}}{\sqrt{\beta_1^2-1}\;\;\sqrt{\beta_2^2-1}}. \tag{21}$$

The integrals in (14) can be calculated if the function $\Omega(w)$ is taken from (21). As a result of the calculations we obtain the following approximate expressions for the reflection and transmission coefficients:

$$R = \frac{\beta_1^2-1}{\beta_1^2(\beta_1+\beta_2)^2}\left(\sqrt{\beta_1^2-1}-\sqrt{\beta_2^2-1}\right)^2 \exp\left(\frac{3}{4a^2k^2}\frac{\sqrt{\beta_2^2-1}-\sqrt{\beta_1^2-1}}{(\beta_1^2-1)\sqrt{\beta_2^2-1}}\right),$$

$$T = \frac{4\sqrt{\beta_1^2-1}\sqrt{\beta_2^2-1}}{\left(\sqrt{\beta_1^2-1}+\sqrt{\beta_2^2-1}\right)^2} \exp\left[\frac{3}{8a^2k^2}\frac{\left(\sqrt{\beta_2^2-1}-\sqrt{\beta_1^2-1}\right)^2}{(\beta_1^2-1)(\beta_2^2-1)}\right]. \tag{22}$$

We note that the numbers preceding the exponential factors in (22) are respectively the reflection and transmission power coefficients R_∞ and T_∞ for a step on an impedance plane. These coefficients can be obtained from the formulas found by Trenev [1]. It is seen from formulas (22) that the dependence of R and T on the radius of the cylinder at large values of a/λ is rather weak.

Figures 1, 2, and 3 show curves of the dependence of R/R_∞ and T/T_∞ on the ratio a/λ, calculated by formulas (20), (14), (16), and (13) by using a digital computer. The fraction of the energy carried by the space wave is best characterized by a radiation coefficient.

$$\eta = 1 - R - T.$$

The curves for η/η_∞ are shown in Figures 4 and 5 (η_∞ is the radiation coefficient for the impedance step on a plane).

It is seen from the curves that an appreciable dependence of R, T, and η on a/λ exists only at small values of a/λ. When $a/\lambda > 1$, the values of R/R_∞, T/T_∞ and η/η_∞ differ little from unity. For such values of a/λ, we can calculate the coefficients R, T, and η from the approximate formulas (22). As $a/\lambda \to 0$, the function X(u) vanishes, and the coefficients R, T, and η tend to the values

$$R_0 = \left(\frac{\beta_1-\beta_2}{\beta_1+\beta_2}\right)^2, \qquad T_0 = \frac{4\beta_1\beta_2}{(\beta_1+\beta_2)^2}, \qquad \eta_0 = 0.$$

Let us calculate now the radiation field resulting from the diffraction of the surface wave on the impedance step. For this purpose we change in the integrals of (17) or (18) to a new variable τ, by means of the substitution

$$w = k\cos\tau, \quad v = k\sin\tau$$

and to a spherical system of coordinates

$$z = r\cos\theta, \quad \rho = z\sin\theta.$$

Estimating the integral in (18) for large values of $k\tau$ by the method of steepest descent (saddle point $\tau = \pi - \theta$), we obtain

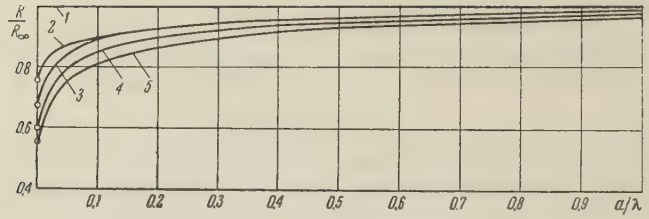


Figure 1. Dependence of the reflection coefficients R/R_∞ on the ratio a/λ ($\beta_2 = 1.05$):

1 — $\beta_1 = 1.05$; 2 — $\beta_1 = 1.1$; 3 — $\beta_1 = 3.0$; 4 — $\beta_1 = 2.0$; 5 — $\beta_1 = 1.5$

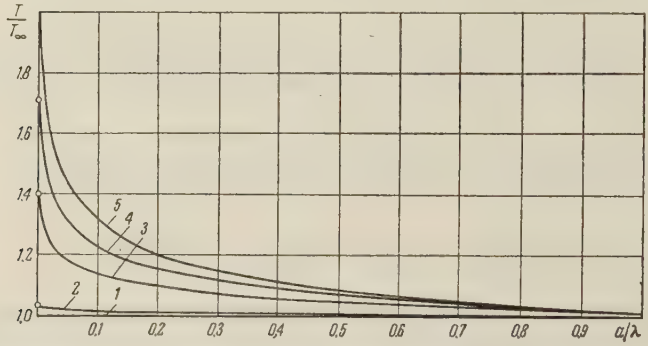


Figure 2. Dependence of the transmission coefficients T/T_∞ on the ratio a/λ ($\beta_2 = 1.05$):

1 — $\beta_1 = 1.05$; 2 — $\beta_1 = 1.1$; 3 — $\beta_1 = 1.5$; 4 — $\beta_1 = 2.0$; 5 — $\beta_1 = 3.0$

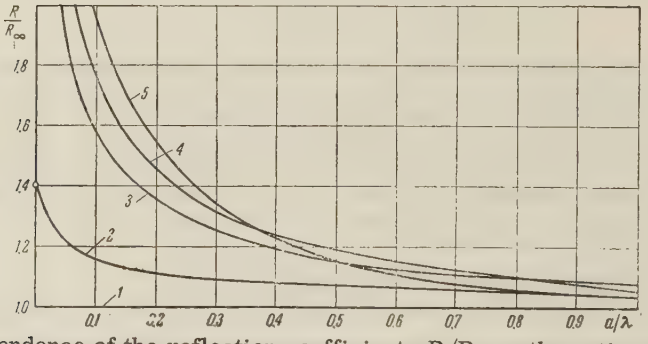


Figure 3. Dependence of the reflection coefficients R/R_∞ on the ratio a/λ ($\beta_1 = 1.05$):

1 — $\beta_2 = 1.05$; 2 — $\beta_2 = 1.1$; 3 — $\beta_2 = 1.5$; 4 — $\beta_2 = 2.0$; 5 — $\beta_2 = 3.0$

$$H(r, \Theta) = \frac{1}{\pi i} \frac{e^{ikr}}{r} \times$$

$$\times \sqrt{2 (Q_2 - Q_1) \frac{h_1^2 (h_1 - h_2)}{k^2 (h_1 + h_2)} ak \{[H_1^{(1)}(av_1)]^2 - H_0^{(1)}(av_1) H_2^{(1)}(av_1)\}} \exp\left[\frac{M(h_1)}{2}\right] \times$$

$$\times \frac{1}{\sqrt{[\sin\Theta H_0^{(1)}(ak\sin\Theta) + Q_1 H_1^{(1)}(ak\sin\Theta)][\sin\Theta H_0^{(1)}(ak\sin\Theta) + Q_2 H_1^{(1)}(ak\sin\Theta)]}} \times$$

$$\times \frac{k(k\cos\Theta + h_2)}{\sqrt{(k^2\cos^2\Theta - h_1^2)(k^2\cos^2\Theta - h_2^2)}} \exp\left[-\frac{M(k\cos\Theta)}{2}\right]. \quad (23$$

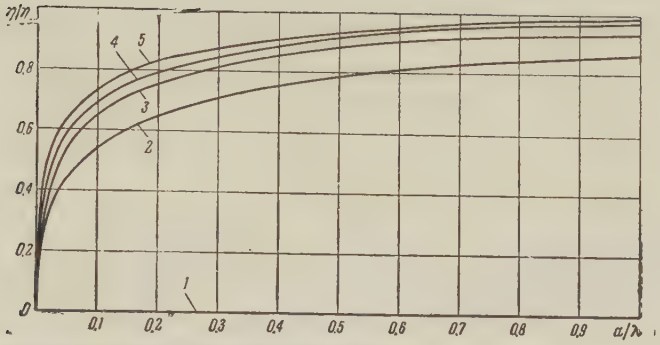


Figure 4. Dependence of the radiation coefficients η/η_∞ on the ratio $a/\lambda(\beta_2 = 1.05)$:

1 — $\beta_1 = 1.05$; 2 — $\beta_1 = 1.1$; 3 — $\beta_1 = 1.5$; 4 — $\beta_1 = 2.0$; 5 — $\beta_1 = 3.0$

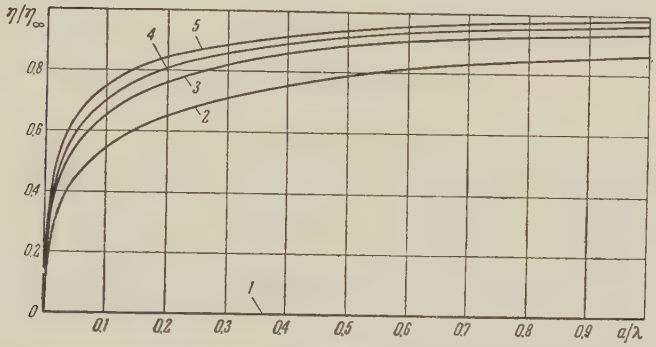


Figure 5. Dependence of the radiation coefficients η/η_∞ on the ratio $a/\lambda(\beta_1 = 1{,}05)$:

1 — $\beta_2 = 1.05$; 2 — $\beta_2 = 1.1$; 3 — $\beta_2 = 1.5$; 4 — $\beta_2 = 2.0$; 5 — $\beta_2 = 3.0$

The power radiation pattern is determined by the function

$$\sum(\theta) = \frac{[\cos\theta + \beta_1) \exp[-X(\cos\theta)]}{(\cos^2\theta - \beta_1^2)(\cos^2\theta - \beta_2^2)\sqrt{(A^2 + B^2)(C^2 + D^2)}}, \tag{24}$$

where

$$A = \sin\theta J_0(ak\sin\theta) K_1(akq_1) + q_1 J_1(ak\sin\theta) K_0(akq_1);$$

$$B = \sin\theta N_0(ak\sin\theta) K_1(akq_1) + q_1 N_1(ak\sin\theta) K_0(akq_1);$$

$$C = \sin\theta J_0(ak\sin\theta) K_1(akq_2) + q_2 J_1(ak\sin\theta) K_0(akq_2);$$

$$D = \sin\theta N_0(ak\sin\theta) K_1(akq_2) + q_2 N_1(ak\sin\theta) K_0(akq_2);$$

$$q_1 = \sqrt{\beta_1^2 - 1}; \qquad q_2 = \sqrt{\beta_2^2 - 1}.$$

The function X(cos θ) in (24) is determined from formula (15).

We do not give the solution for the TE wave, since all the formulas obtained for the TM wave can be modified for the TE wave by replacing H_φ by E_φ, E_ψ by $-H_\rho$, and E_Z by $-H_Z$, and Z by 1/Z.

REFERENCES

1. N.G. Trenev, Diffraction of surface electromagnetic waves on an impedance step. Radiotekhnika i Elektroniak 1958, 3, 1, 27.
2. G. Weill, Etude d'un probleme de diffraction des ondes electromagnetiques de surface application a la theorie de l'antenna dielectrique, Ann. radioelectr., 1955, 10, 41, 228.
3. A.F. Kay, Scattering of a surface wave by a discontinuity on reactance IRE trans., 1959, AP-7, 1, 22.
4. V.I. Talanov, Diffraction of electromagnetic waves on a surface-impedance step in a waveguide, Izvestiya Vuzov, Ministry of Higher Education USSR — Radiophysics, 1957, 1, 3, 64.
5. V.A Fok, Certain Integral Equations of Mathematical Physics, Mathematics Collection, New Series, 1944, 14, 1 — 2, 3.

Submitted to the editors 18 January 1961

THE DETERMINATION OF THE NATURAL FREQUENCIES OF THE HIGHER MODES IN A COAXIAL CABLE WITH ELLIPTICAL CONDUCTORS

V.Ya. Smorgonskiy

The eigenvalues of the transcendental equation are determined for three modes (E_{01}, and two H_{11} modes) in a ring between two confocal ellipses as a function of the eccentricity of the external ellipse, for two values of the ratio of the major axes of the two ellipses.

The purpose of the work was a calculation of the effect of eccentricity of the conductors of a coaxial cable on the cutoff frequency of the higher modes. The influence of the conductor eccentricity on the electrical parameters of a coaxial cable was examined in [1—3]. In [1] and [2] the question of higher modes is not discussed. Only in [3] is a general formula given for the cutoff frequency of the E_{01} mode. However, the theoretical formula [Equation (24) of page 359] yields an expression which is independent of the eccentricity, since the approximation used by the author is such that Equation (19) determines the cutoff frequency of the E_{01} mode in a coaxial cable with round conductors (see for example, [4], page 527), that is, the formula given does not make it possible to estimate the effect of the eccentricity of the conductors on the cutoff frequency. Since the lowest value of the cutoff frequency is that of the H_{11} mode, greatest interest attaches to a determination of the effect of eccentricity of conductors on the cutoff frequency at this mode. The eigenfunction of an elliptical ring can be represented in the following form [3]:

$$\Psi = [Je,\ o_n(\xi,\ h) + ANe,\ o_n(\xi,\ h)]\, se,\ o_n(\eta,\ h). \tag{1}$$

Here A is a coefficient determined by the boundary conditions; se_n and so_n are even and odd angular Mathieu functions and Je_n, Jo_n, Ne_n, and No_n are radial (modified) Mathieu

functions of the first and second order n;*h is a parameter determined from the transcendental equations, which in the case when the internal and external conductors are of infinite conductivity have the following form:

$$\frac{Je,\ o_n(\xi_1, h)}{Je,\ o_n(\xi_2, h)} = \frac{Ne,\ o_n(\xi_1, h)}{Ne,\ o_n(\xi_2, h)} \tag{2a}$$

for TM waves and

$$\frac{J'e,\ o_n(\xi_1, h)}{J'e,\ o_n(\xi_2, h)} = \frac{N'e,\ o_n(\xi_1, h)}{N'e,\ o_n(\xi_2, h)} \tag{2b}$$

for TE waves. ξ_1 and ξ_2 are the coordinates of the ellipses corresponding to the surfaces of the internal and external conductors. $\xi_1 = \cosh^{-1}(1/e_1)$; $\xi_2 = \cosh^{-1}(1/e_2)$; e_1 and e_2 are the eccentricities of the conductors. The prime denotes differentiation with respect to ξ. The cutoff frequency is connected with the parameter h, determined from Equations (2a) and (2b), by the following relations [5]:

$$\omega_{cr} = \frac{h}{a_1 e_1 \sqrt{\varepsilon\mu}}.$$

Since we consider a cable whose conductors are confocal ellipses with major semiaxes a_1 and a_2, when the eccentricity of the outer conductor is $e_2 = a_1/a_2$, the inner conductor turns into an infinitesimally thin strip $2a_1$ wide, the equation of which is $\xi = 0$ when $e_2 = 1/k$. Here $k = a_2/a_1$.

In this case the transcendental equations become

$$\frac{Je,\ o_n(o, h)}{Je,\ o_n(\xi_2, h)} = \frac{Ne,\ o_n(o, h)}{Ne,\ o_n(\xi_2, h)} \tag{3a}$$

for the TM modes and

$$\frac{J'e,\ o_n(o, h)}{J'e,\ o_n(\xi_2, h)} = \frac{N'e,\ o_n(o, h)}{N'e,\ o_n(\xi_2, h)} \tag{3b}$$

for the TE modes.

In determining the effect of eccentricity on the value of the cutoff frequency of the H_{11} mode, it is necessary to bear in mind that the cutoff frequencies of even eH_{11} and odd oH_{11} are different. For the even wave with $1/k < e_2 < 1$ the strip $\xi = 0$, to which the internal conductor degenerates, is perpendicular to the electric field lines, that is, the structure of the field is the same as in a hollow elliptic waveguide, and the cutoff frequency is determined from the expression

$$J'e_1(\xi_2, h) = 0. \tag{4}$$

For the odd wave condition, (3b) with $1/k < e_2 < 1$ will be valid for an internal-strip length $a_1 = a_2 e_2$. To determine the values of h satisfying equations (2) and (3) it is convenient to use the representation of the radial Mathieu functions in the form of rapidly converging

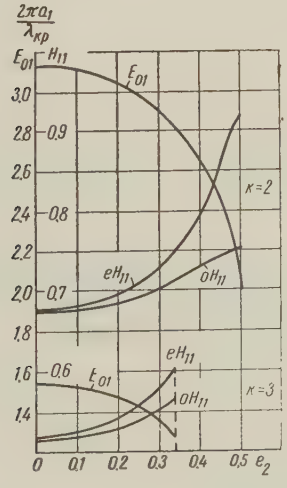


Figure 1. Values of cutoff wave numbers for eH_{11}, $0H_{11}$, and E_{01} modes in a coaxial cable with elliptic conductors.

series of products of cylindrical functions ([6], page XXXI)

*We use in the text the notation of [5] for the Mathieu functions.

The results of the calculations are shown in the figure, where the eigenvalues of the corresponding transcendental equations are shown as functions of the eccentricity of the external conductor. The lowest cutoff frequency is possessed by the OH_{11} mode. The cutoff frequency of the E_{01} mode is higher than the cutoff frequencies of the eH_{11} and $0H_{11}$ mode.

The relative change in cutoff frequencies with increasing eccentricity is faster at large values of k; thus, for example, for the eH_{11} mode with $e_2 = 0.3$ the relative change in cutoff frequency is 9 percent when k = 2 and 12 percent when k = 3. At an external-conductor eccentricity $e_2 < 0.1$, the value of the cutoff frequencies remains almost constant ($\Delta f/f < 0.01$). The rapid change in cutoff frequencies begins with $e_2 > 0.2$.

REFERENCES

1. M.R. Shebes, Radiotekhnika, 1949, 4, 4, 36.
2. S. Mahaparta, Electronic and Radio Engr., 1958, 35, 2, 63.
3. J. Wong, Canad, J. Phys, 1956, 34, 4, 354.
4. A.D. Gol'dshteyn, N.V. Zernov, Electromagnetic Fields and Waves, Sovetskoye Radio, 1956.
5. P.M. Morse and H. Feshbach, Methods of Theoretical Physics, McGraw Hill, 1956.
6. Tables relating to Mathieux functions, Cambridge University Press, N.Y. 1951.

Submitted to the editors 1 December 1960

CONCERNING A NONLINEAR PROBLEM IN THE THEORY OF ELECTROMAGNETIC WAVES IN A PLASMA

Yu. N. Dnestrovskiy, D.P. Kostomarov

It is shown that for an electromagnetic wave propagating in a plasma transversely to an external magnetic field H_0, and polarized along this field (ordinary wave), the results of the nonlinear theory and the theory using linearization of the transport equation are the same.

We consider electromagnetic waves in a magnetoactive plasma, propagating perpendicular to the external magnetic field H_0. We shall assume the field H_0 homogeneous, and the electromagnetic vector polarized along this field (ordinary wave). Let the x-axis be parallel to the propagation of the wave, and the z-axis parallel to the field H_0.

The process is described by the following system of equations

$$\frac{\partial f}{\partial t} + v\cos\delta\frac{\partial f}{\partial x} - \omega_H\frac{\partial f}{\partial \delta} + \frac{e}{m}E(t, x)\frac{\partial f}{\partial u} = 0, \tag{1}$$

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial j}{\partial t} = \frac{4\pi e}{c^2}\frac{\partial}{\partial t}\int_0^{2\pi} d\delta\int_0^{\infty} vdv\int_{-\infty}^{\infty} uf\,du, \tag{2}$$

where v, δ, and u are cylindrical coordinates in velocity space; $f = f(t, x, v, \delta, u)$ — the electron distribution function; $\omega_H = eH_0/mc$ — the Larmor frequency; $E(t,x) = E_Z(t,x)$; $j(t,x) = j_Z(t,x)$.

The effect of the magnetic field of the wave on the electrons is neglected.

The first integrals of the characteristic system of equation (1) have the form

$$g_1 = \delta + \omega_H t = \text{const}, \quad g_2 = x + \frac{v}{\omega_H}\sin\delta = \text{const},$$

$$g_3 = u - \frac{e}{m}\int_{t_*}^{t} E\left[\tau, x + \frac{v}{\omega_H}\sin\delta - \frac{v}{\omega_H}\sin(\delta + \omega_H t - \omega_H\tau)\right]d\tau = \text{const}.$$

The general solution of (1) is an arbitrary function of g_1, g_2, and g_3 and of the variable v, which enters into this equation as a parameter. We take for the distribution function f such solutions of equation (1) which are independent of g_1:

$$f = f(v, g_2, g_3). \tag{3}$$

Here f is an arbitrary nonnegative function of its arguments, integrable within infinite limits in velocity space and even with respect to g_3. The distribution function depends on time only through the electric field. When $E(t,x) = 0$, the distribution becomes stationary; $f = f(v, x + (v/\omega_H)\sin\delta, u)$, and the fact that the function is even with respect to the third argument denotes the equivalence of the positive and negative directions of the z-axis in the absence of an electric field.

If the distribution function is explicitly dependent on g_2, the plasma in inhomogeneous. In this case, as can be seen from (3), there exists in the plasma a stationary current, parallel to the y-axis;

$$j_y(x) = e\int_0^{2\pi}\sin\delta d\delta\int_0^{\infty} v^2 dv\int_{-\infty}^{\infty} f\left(v, x + \frac{v}{\omega_H}\sin\delta, u\right)du.$$

The current produced an additional inhomogeneous stationary magnetic field in the z-direction, the pressure of which should be compensated by the inhomogeneity of the pressure of the plasma. We shall disregard the inhomogeneity of the magnetic field. This approximation is valid if the quantity μ_0, the ratio of the plasma pressure to the magnetic pressure, is small. Then the error that enters into the determination of the electric wave field will be of the order μ_0^2 [1].

Let us return to the system (1) — (2). We calculate the current in the z-direction:

$$j(t,x) = e\int_0^{2\pi} d\delta\int_0^{\infty} vdv\int_{-\infty}^{\infty} f(v, g_2, g_3)\,udu =$$
$$= \frac{e^2}{m}\int_0^{2\pi} d\delta\int_0^{\infty}\varphi\left(v, x + \frac{v}{\omega_H}\sin\delta\right)vdv\int_{t_*}^{t} E\left[\tau, x + \frac{v}{\omega_H}\sin\delta - \right.$$
$$\left. - \frac{v}{\omega_H}\sin(\delta + \omega_H t - \omega_H\tau)\right]d\tau. \tag{4}$$

Here

$$\varphi\left(v, x + \frac{v}{\omega_H}\sin\delta\right) = \int_{-\infty}^{\infty} f\left(v, x + \frac{v}{\omega_H}\sin\delta, u\right)du. \tag{5}$$

Substituting (4) into (2) we obtain a linear integral-differential equation with respect to the field $E(t,x)$:

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = \frac{4\pi e^2}{mc^2}\frac{\partial}{\partial t}\int_0^{2\pi} d\delta\int_0^{\infty}\varphi\left(v, x + \frac{v}{\omega_H}\sin\delta\right)vdv \times$$
$$\times\int_{t_*}^{t} E\left[\tau, x + \frac{v}{\omega_H}\sin\delta - \frac{v}{\omega_H}\sin(\delta + \omega_H t - \omega_H\tau)\right]d\tau. \tag{6}$$

We emphasize that Equation (6) agrees fully with the equation for the electric field obtained by the usual procedure of linearizing Equation (1), but our derivation of this equation does not involve the assumption that the change produced in the function f by the wave field is small.

We consider now fields that have a harmonic dependence on the time:

$$E(t, x) = e^{-i\omega t}E(x) \qquad \left(\frac{\omega}{c} = k\right). \tag{7}$$

We put in (4) and (6) $t_0 = -\infty$ and consider the internal integral with respect to τ

$$I = \int_{-\infty}^{t} e^{-i\omega\tau}E\left[x + \frac{v}{\omega_H}\sin\delta - \frac{v}{\omega_H}\sin(\delta + \omega_H t - \omega_H\tau)\right]d\tau.$$

Introducing a new integration variable $z = \delta + \omega_H t - \omega_H\tau$ and carrying out some simple transformations, we obtain

$$I = e^{-i\omega t}\frac{1}{2\omega_H \sin\frac{\omega}{\omega_H}\pi}\int_{\delta-2\pi}^{\delta} e^{i\frac{\omega}{\omega_H}(\alpha-\delta+\pi)}E\left[x + \frac{v}{\omega_H}(\sin\delta - \sin\alpha)\right]d\alpha.$$

Thus, for single-frequency waves, Equation (6) assumes the form

$$\frac{d^2E}{dx^2} + k^2E = \frac{2\pi e^2}{mc^2}\frac{\frac{\omega}{\omega_H}}{\sin\frac{\omega}{\omega_H}\pi}\int_0^{2\pi}d\delta\int_{\delta-2\pi}^{\delta} e^{i\frac{\omega}{\omega_H}(\alpha-\delta+\pi)}d\alpha \times$$

$$\times\int_0^{\infty}\varphi\left(v, x + \frac{v}{\omega_H}\sin\delta\right)E\left[x + \frac{v}{\omega_H}(\sin\delta - \sin\alpha)\right]v\,dv. \tag{8}$$

This equation, too, naturally coincides with the corresponding equation of the linearized theory.

The first part of Equation (8) becomes infinite when the frequency ω is a multiple of the Larmor frequency ω_H. In this case the motion of the electrons over the Larmor circles becomes resonant with the oscillations in the field of the wave. In the vicinity of the resonant frequencies the effect of the electric wave field on the distribution function of the electrons is no longer small and consequently the ordinary approach to the solution of the problem, based on the linearization of the transport equation, is not valid. The fact that Equation (8) is derived from the nonlinear system (1) and (2) indicates that it is valid (within the framework of the approximation considered) also for frequencies close to resonance. The question of how unaccounted for effects will influence the vicinity of the resonant frequencies (effect of the magnetic field of the wave, collisions of particles, etc.) necessitates a special investigation.

In conclusion we note that there exists still another case when the problem leads in the nonlinear formulation to the same equation for the field as the linear problem [2]. It is typical that in both cases the wave is purely transverse If the electric vector has a longitudinal component, the equations for the field becomes essentially nonlinear. In this case the linearization procedure will naturally lead to a different equation.

REFERENCES

1. Yu. N. Dnestrovskiy, E.P. Kostomarov, Zh. Eksp. Teoret, Fiz. 1960, 39, 9, 845.
2. R.Z. Sagdeyev, Plasma Physics and the Problem of Controllable Thermonuclear Reactions (Collection), Izd. AN SSSR, 1958, Vol. 4, 422.

Physics Faculty
of the M.V. Lomonosov
Moscow State University

Submitted to the editors, 22 February 1961

CONTRIBUTION TO THE DESIGN OF RECTANGULAR WAVEGUIDES LOADED WITH DIELECTRIC DIAPHRAGMS

A.N. Didenko and L.N. Bezmaternykh

A dispersion equation is derived for the inphase waves of a rectangular waveguide, loaded on the sidewalls with dielectric diaphragms. Along with the rigorous solution, an approximate dispersion equation is obtained, which permits a simpler account of the higher spatial harmonics.

Recently the foreign and the domestic literature have contained detailed analysis of the effectiveness and dispersion properties of cylindrical waveguides loaded with dielectric diaphragms [1, 2]. It has been established theoretically and experimentally that when a dielectric with a large value of ϵ is used and when the diaphragm thickness is suitably chosen, it is possible to obtain slow-wave systems with shunt resistances which are at least one order of magnitude higher than the shunt resistances of waveguides with metallic diaphragms.

Slow-wave systems with high shunt resistances are usually used in high-power traveling wave tubes and in linear electron accelerators of high energy, where the question of reducing the power of high-frequency oscillators is of prime significance. In addition, closed, bent diaphragmed waveguides can be used as accelerating devices for electronic cyclic accelerators of high energy [3].

If the technology of manufacture of the waveguides is on a sufficiently high level, the effectiveness of such systems can be one order of magnitude greater than the effectiveness of ordinary resonator systems. The analysis performed has shown that in this case the most suitable are bent rectangular waveguides in which the diaphragms are located along the bent walls, and that of all the modes existing in such waveguides, the most suitable is the symmetrical LE_{11} mode, which has on the mean radius a large azimuthal component of electric field.

Reference [3] considers unbent rectangular waveguides with metallic diaphragms, since for large synchrotrons the bending will be slight and therefore in the investigation of various electrodynamic characteristics of such slow-wave systems we can neglect the bending in the first approximation.

In the present paper we develop the fundamentals of the theory of waveguides of rectangular cross section, loaded with dielectric diaphragms.

The system under consideration is a rectangular waveguide of height a and width b, inside of which, on the side walls, are located dielectric diaphragms of thickness L_2 and height h, spaced L_1 apart; the diaphragms are made of an isotropic dielectric with large value of ϵ. Following [2], we derive the dispersion equations of this system by starting from the dispersion equations for the waves existing in this system in the absence of central openings in the dielectric. It will be shown below that if the fields depend on both transverse coordinates, then only hybrid waves can exist in this system, and it is therefore necessary to obtain simultaneous dispersion equations for the TE and TM modes which exist in waves with solid diaphragms.

Writing down the expressions for the T^E and TM modes and equating their tangential components on the air-dielectric interface, we readily obtain the following dispersion equations:

$$\cos k_3 L = \cos p_1 L_1 \cos p_2 L_2 - \frac{\varepsilon^2 p_1^2 + p_2^2}{2\varepsilon p_1 p_2} \sin p_1 L_1 \sin p_2 L_2 \quad (TM), \tag{1}$$

$$\cos k_3 L = \cos p_1 L_1 \cos p_2 L_2 - \frac{p_1^2 + p_2^2}{2 p_1 p_2} \sin p_1 L_1 \sin p_2 L_2 \quad (TE), \tag{2}$$

where

$$k_3 = \frac{\omega}{v_\Phi} = \frac{2\pi}{\lambda_g};$$
$$L = L_1 + L_2;$$
$$p_1^2 = k^2 - \Gamma^2 - \left(\frac{\pi}{a}\right)^2;$$
$$p_2^2 = \varepsilon k^2 - \Gamma^2 - \left(\frac{\pi}{a}\right)^2.$$

These equations determine the dispersion properties of our system when the diaphragms are solid. In the case of thin and closely spaced diaphragms ($p_1 L_1 \ll 1$, $p_2 L_2 \ll 1$) the equations reduce to the following

$$\left(\frac{\pi}{a}\right)^2 + (\Gamma^e)^2 = \frac{\varepsilon(L_1+L_2)}{\varepsilon L_1 + L_2} k^2 - \frac{\varepsilon(L_1+L_2)^2}{(L_1+\varepsilon L_2)(\varepsilon L_1 + L_2)} k_3^2, \tag{3}$$

$$\left(\frac{\pi}{a}\right)^2 + (\Gamma^m)^2 = \frac{\varepsilon(L_1+L_2)}{\varepsilon L_1 + L_2} k^2 - \frac{L_1+L_2}{L_1+\varepsilon L_2} k_3^2, \tag{4}$$

where Γ^e and Γ^m are the transverse wave numbers for the TM and TE modes in the y-direction.

It is seen from (3) and (4) that in this case a system without openings is equivalent to a system filled with an anisotropic dielectric with $\epsilon_x = \epsilon_y = (L_1 + \epsilon L_2)/(L_1 + L_2)$ and $\epsilon_z = \epsilon(L_1 + L_2)/\epsilon(L_1 + L_2)$. In the general case, however, the connection between the components of the wave vector can be determined by solving the transcendental equations (1) and (2). We shall refer to this in greater detail later on.

The field components for T^M and TE modes can be obtained analogously, and will therefore not be written out here.

We now consider a waveguide in which the dielectric diaphragms have central openings for passage of a beam. The presence of openings makes it impossible to satisfy the boundary conditions if only the TE or the TM modes are taken into account separately. This means that in such a system only hybrid waves can exist. Using the fact that the field components of a hybrid wave can be expressed in terms of the electric and magnetic Hertz vectors.

$$\vec{E} = \text{grad div } \vec{\Pi}_e - \Delta\vec{\Pi}_e + ik \text{ rot } \vec{\Pi}_m, \tag{5}$$

$$\vec{H} = -i\varepsilon k \text{ rot } \vec{\Pi}_e + \text{grad div } \vec{\Pi}_m - \Delta\vec{\Pi}_m \tag{6}$$

and assuming as usual that the vectors $\vec{\Pi}_e$ and $\vec{\Pi}_m$ have only a z-component, we obtain finally the following expressions for the components of the inphase hybrid wave in the interaction space $\left(0 \leqslant |y| \leqslant \left|\frac{b}{2} - h\right|\right)$:

$$\begin{aligned}
E_x &= \sin\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} \text{ch } k_{y,s}y\,(-i\beta_s Q_s^{(1)} A_s + Q_s^{(2)} B_s)\, e^{i\beta_s z},\\
E_y &= \cos\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} \text{sh } k_{y,s}y\,(i\beta_s Q_s^{(3)} A_s - Q_s^{(4)} B_s)\, e^{i\beta_s z},\\
E_z &= \cos\frac{\pi}{2}x \sum_{s=-\infty}^{\infty} \text{ch } k_{y,s}y A_s e^{i\beta_s z},\\
H_x &= \cos\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} \text{sh } k_{y,s}y\,(-Q_s^{(2)} A_s + i\beta_s Q_s^{(1)} B_s)\, e^{i\beta_s z},\\
H_y &= \sin\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} \text{ch } k_{y,s}y\,(-Q_s^{(4)} A_s + i\beta_s Q_s^{(3)} B_s)\, e^{i\beta_s z},\\
H_z &= \sin\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} \text{sh } k_{y,s}y\, B_s e^{i\beta_s z}.
\end{aligned} \tag{7}$$

Here

$$Q_s^{(1)} = \frac{\frac{\pi}{a}}{\left(\frac{\pi}{a}\right)^2 - k_{y,s}^2}; \qquad Q_s^{(2)} = \frac{ikk_{y,s}}{\left(\frac{\pi}{a}\right)^2 - k_{y,s}^2};$$

$$Q_s^{(3)} = \frac{k_{y,s}}{\left(\frac{\pi}{a}\right)^2 - k_{y,s}^2}; \qquad Q_s^{(4)} = \frac{ik\frac{\pi}{a}}{\left(\frac{\pi}{a}\right)^2 - k_{y,s}^2}; \tag{8}$$

$$-k_{y,s}^2 = k^2 - \beta_s^2 - \left(\frac{\pi}{a}\right)^2;$$

β_S is the propagation constant of the s-th spatial harmonic, connected with the fundamental harmonic by the relation

$$\beta_s = \beta_0 + \frac{2\pi s}{L}.$$

For the second region $\left(\left|\frac{b}{2} - h\right| \leqslant |y| \leqslant \left|\frac{b}{2}\right|\right)$, using the expressions for the components of the TE and TM waves in a system without apertures, we can readily obtain the following expression for the field components of the hybrid wave.

$$E_x = \sin\frac{\pi}{a}x \sum_{r=-\infty}^{\infty}\left[-R_r^{e(1)}a_r \operatorname{sh}\Gamma_r^e\left(y-\frac{b}{2}\right)\frac{dW_r^e}{dz} + R_r^{m(2)}b_r \operatorname{sh}\Gamma_r^m\left(y-\frac{b}{2}\right)W_r^m\right],$$

$$E_y = \cos\frac{\pi}{a}x \sum_{r=-\infty}^{\infty}\left[R_r^{e(3)}a_r \operatorname{ch}\Gamma_r^e\left(y-\frac{b}{2}\right)\frac{dW_r^e}{dz} - R_r^{m(4)}b_r \operatorname{ch}\Gamma_r^m\left(y-\frac{b}{2}\right)W_r^m\right],$$

$$E_z = \cos\frac{\pi}{a}x \sum_{r=-\infty}^{\infty} a_r \operatorname{sh}\Gamma_r^e\left(y-\frac{b}{2}\right)W_r^e, \tag{9}$$

$$H_x = \cos\frac{\pi}{a}x \sum_{r=-\infty}^{\infty}\left[-\varepsilon R_r^{e(2)}a_r \operatorname{ch}\Gamma_r^e\left(y-\frac{b}{2}\right)W_r^e + R_r^{m(1)}b_r \operatorname{ch}\Gamma_r^m\left(y-\frac{b}{2}\right)\frac{dW_r^m}{dz}\right],$$

$$H_y = \sin\frac{\pi}{a}x \sum_{r=-\infty}^{\infty}\left[-\varepsilon R_r^{e(4)}a_r \operatorname{sh}\Gamma_r^e\left(y-\frac{b}{2}\right)W_r^e + R_r^{m(3)}b_r \operatorname{sh}\Gamma_r^m\left(y-\frac{b}{2}\right)\frac{dW_r^m}{dz}\right],$$

$$H_z = \sin\frac{\pi}{a}x \sum_{r=-\infty}^{\infty} b_r \operatorname{ch}\Gamma_r^m\left(y-\frac{b}{2}\right)W_r^m,$$

where Γ_r^e and Γ_r^m are the roots of the equations for the TE and TM modes for specified k_3;

$$R_r^{e(1)} = \frac{\frac{\pi}{a}}{\left(\frac{\pi}{a}\right)^2 - (\Gamma_r^e)^2}; \qquad R_r^{e(2)} = \frac{ik\Gamma_r^e}{\left(\frac{\pi}{a}\right)^2 - (\Gamma_r^e)^2};$$

$$R_r^{e(3)} = \frac{\Gamma_r^e}{\left(\frac{\pi}{a}\right)^2 - (\Gamma_r^e)^2}; \qquad R_r^{e(4)} = \frac{ik\frac{\pi}{a}}{\left(\frac{\pi}{a}\right)^2 - (\Gamma_r^e)^2} \tag{10}$$

(The expressions for $R_r^{m(1)}$ and $R_r^{m(4)}$ can be obtained by replacing Γ_r^e by Γ_r^m);

$$W^e(z) = u_1^e(z) - \frac{u_1^e(L) - e^{i\psi}}{u_2^e(L)}u_2^e(z); \tag{11}$$

The functions u_1^e and u_2^e are solutions of the equation

$$\frac{d^2u^e(z)}{dz^2} + \left[\varepsilon k^2 - \left(\frac{\pi}{a}\right)^2 - (\Gamma^e)^2\right] u^e(z) = 0. \tag{12}$$

When $\epsilon = 1$ for $0 \leqslant z \leqslant L_1$ and $\epsilon = \epsilon$ for $L_1 \leqslant z \leqslant L_1 + L_2$, these functions satisfy the initial conditions

$$u_1^e(0) = 1; \quad \frac{1}{\varepsilon}\frac{du_1^e(0)}{dz} = 0; \quad u_2^e(0) = 0; \frac{1}{\varepsilon}\frac{du_2^e(0)}{dz} = 1 \tag{13}$$

and have the following form

$$u_1^e(z) = \cos p_1 z \quad \text{for } 0 \leqslant z \leqslant L_1,$$

$$u_1^e(z) = \frac{1}{\varepsilon}\cos p_1 L_1 \cos p_2 (z - L_1) - \frac{p_1}{p_2}\sin p_1 L_1 \sin p_2 (z - L_1)$$
$$\text{for } L_1 \leqslant z \leqslant L_1 + L_2; \tag{14}$$

$$u_e^2(z) = \frac{1}{p_1}\sin p_1 z \quad \text{for } 0 \leqslant z \leqslant L_1,$$

$$u_2^e(z) = \frac{1}{\varepsilon p_1}\sin p_1 L_1 \cos p_2 (z - L_1) + \frac{1}{p_2}\cos p_1 L_1 \sin p_2 (z - L_1) \tag{15}$$
$$\text{for } L_1 \leqslant z \leqslant L_1 + L_2$$

(the formulas for u_1^m and u_2^M can be obtained be eliminating p_2 and putting $\epsilon = 1$ everywhere in the formulas for u_1^e and u_2^e).

Expanding W_r^e and W_r^m and their derivatives in spatial harmonics we obtain finally the following expressions for the field components in the second region

$$E_x = \sin\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} e^{i\beta_s z} \sum_{r=-\infty}^{\infty} \left[-R_r^{e(1)} a_r \operatorname{sh}\Gamma_r^e\left(y - \frac{b}{2}\right) T_{r,s}^e + \right.$$
$$\left. + R_r^{m(2)} b_r \operatorname{sh}\Gamma_r^m\left(y - \frac{b}{2}\right) S_{r,s}^m\right],$$

$$E_y = \cos\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} e^{i\beta_s z} \sum_{r=-\infty}^{\infty} \left[R_r^{e(3)} a_r \operatorname{ch}\Gamma_r^e\left(y - \frac{b}{2}\right) T_{r,s}^e - \right.$$
$$\left. - R_r^{m(4)} b_r \operatorname{ch}\Gamma_r^m\left(y - \frac{b}{2}\right) S_{r,s}^m\right],$$

$$E_z = \cos\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} e^{i\beta_s z} \sum_{r=-\infty}^{\infty} a_r \operatorname{sh}\Gamma_r^e\left(y - \frac{b}{2}\right) S_{r,s}^e,$$

$$H_x = \cos\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} e^{i\beta_s z} \sum_{r=-\infty}^{\infty} \left[-R_r^{e(2)} a_r \operatorname{ch}\Gamma_r^e\left(y - \frac{b}{2}\right) \tilde{S}_{r,s}^e + \right. \tag{16}$$
$$\left. + R_r^{m(1)} b_r \operatorname{ch}\Gamma_r^m\left(y - \frac{b}{2}\right) T_{r,s}^m\right],$$

$$H_y = \sin\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} e^{i\beta_s z} \sum_{r=-\infty}^{\infty} \left[-R_r^{e(4)} a_r \operatorname{sh}\Gamma_s^e\left(y - \frac{b}{2}\right) \tilde{S}_{r,s}^e + \right.$$
$$\left. + R_r^{m(3)} b_r \operatorname{sh}\Gamma_r^m\left(y - \frac{b}{2}\right) T_{r,s}^m\right],$$

$$H_z = \sin\frac{\pi}{a}x \sum_{s=-\infty}^{\infty} e^{i\beta_s z} \sum_{r=-\infty}^{\infty} b_r \operatorname{ch}\Gamma_r^m\left(y - \frac{b}{2}\right) S_{r,s}^m,$$

where

$$S^e_{r,s} = \frac{1}{L}\int_0^L e^{-i\beta_s z} W^e_r(z)\,dz = \frac{1}{L}(S^{e\,(1)}_{r,s} + S^{e\,(2)}_{r,s}); \tag{17}$$

$$\tilde{S}^e_{r,s} = \frac{1}{L}\int_0^L \varepsilon e^{-i\beta_s z} W^e_r(z)\,dz = \frac{1}{L}(S^{e\,(1)}_{r,s} + \varepsilon S^{e\,(2)}_{r,s}); \tag{18}$$

$$T^e_{r,s} = \frac{1}{L}\int_0^L e^{-i\beta_s z}\frac{dW^e_r(z)}{dz}\,dz = \frac{1}{L}(T^{e\,(1)}_{r,s} + T^{e\,(2)}_{r,s}); \tag{19}$$

$$S^{e\,(1)}_{r,s} = -\frac{1}{p_1^2-\beta_s^2}\left\{e^{-i\beta_s L_1}\left[i\beta_s W^e_r(L_1) + \frac{dW^e_r(L_1)}{dz}\right] - \left[i\beta_s W^e_r(0) + \frac{dW^e_r(0)}{dz}\right]\right\}; \tag{20}$$

$$S^{e\,(2)}_{r,s} = -\frac{1}{p_2^2-\beta_s^2}\left\{e^{-i\beta_s L}\left[i\beta_s W^e_r(L) + \frac{dW^e_r(L)}{dz}\right] - e^{-i\beta_s L_1}\left[i\beta_s W^e_r(L_1) + \frac{dW^e_r(L_1)}{dz}\right]\right\}; \tag{21}$$

$$T^{e\,(1)}_{r,s} = -\frac{1}{p_1^2-\beta_s^2}\left\{e^{-i\beta_s L_1}\left[i\beta_s \frac{dW^e_r(L_1)}{dz} - p_1^2 W^e_r(L_1)\right] - \right. \tag{22}$$

$$\left. -\left[i\beta_s \frac{dW^e_r(0)}{dz} - p_1^2 W^e_r(0)\right]\right\};$$

$$T^{e\,(2)}_{r,s} = -\frac{1}{p_2^2-\beta_s^2}\left\{e^{-i\beta_s L}\left[i\beta_s \frac{dW^e_r(L)}{dz} - p_1^2 W^e_r(L)\right] - \right. \tag{23}$$

$$\left. - e^{-i\beta_s L_1}\left[i\beta_s \frac{dW^e_r(L_1)}{dz} - p_1^2 W^e_r(L_1)\right]\right\}.$$

The expressions for $S^m_{r,s}$ and $T^m_{r,s}$ can be obtained by replacing in (17) - (23) W^e_r and dW^e_r/dz by W^m_r and dW^m_r/dz.

Equating the corresponding harmonics of the tangential components of the field (7) and (16) with $y = y_0 = (b/2) - h$ and eliminating A_s and B_s, we obtain the following two infinite systems of equations linear in a_r and b_r, which determine the dispersion properties of the inphase waves of a rectangular waveguide with dielectric diaphragms

$$\sum_{r=-\infty}^{\infty} a_r\,[Q^{(2)}_s S^e_{r,s}\,\mathrm{th}\,k_{y,s}y_0\,\mathrm{sh}\,\Gamma^e_r h + R^{e\,(2)}_r \tilde{S}^e_{r,s}\,\mathrm{ch}\,\Gamma^e_r h] + \tag{24}$$

$$+ b_r\,[i\beta_s Q^{(1)}_s S^m_{r,s}\,\mathrm{ch}\,\Gamma^m_r h - R^{m\,(1)}_r T^m_{r,s}\,\mathrm{ch}\,\Gamma^m_r h] = 0,$$

$$\sum_{r=-\infty}^{\infty} a_r\,[i\beta_s Q^{(1)}_s S^e_{r,s}\,\mathrm{sh}\,\Gamma^e_r h - R^{e\,(1)}_r T^e_{r,s}\,\mathrm{sh}\,\Gamma^e_r h] +$$

$$+ b_r\,[Q^{(2)}_s S^m_{r,s}\,\mathrm{cth}\,k_{y,s}y_0\,\mathrm{ch}\,\Gamma^m_r h + R^{m\,(2)}_r S^m_{r,s}\,\mathrm{sh}\,\Gamma^m_r h] = 0.$$

Here s runs through all values from $-\infty$ to $+\infty$.

The system has a solution if its determinant vanishes. The order of its determinant in this case increases very rapidly (if the first harmonics in the forward and backward direction are taken, we have a sixth-order determinant), and it is therefore quite difficult to investigate the dispersion properties of this slow-wave system by starting we assume that the field distribution in the second region $\left(\left|\frac{b}{2}-h\right| \leqslant |y_1| \leqslant \left|\frac{b}{2}\right|\right)$ is described with sufficient accuracy by the functions W^e_0 and W^m_0 and their derivatives (that is, if we disregard the summation over r) , then the higher spatial harmonics can be taken into account with the aid of the approximate method of joining the partial powers, as proposed by L.A. Vaynshteyn [4] . The gist of this method is to replace the exact joining of the values, which in this case is quite difficult by calculation of the partial powers for regions I and II over the periods L of the structure, and joining them on the separation boundary;

$$p_1^{\mathrm{I}} = p_1^{\mathrm{II}},\quad p_2^{\mathrm{I}} = p_2^{\mathrm{II}},$$

where

$$p_1 = \frac{c}{8\pi}\int_0^L E_z^* H_x dz; \qquad (25)$$

$$p_2 = \frac{c}{8\pi}\int_0^L E_x^* H_z\, dz.$$

Having done so, we finally obtain the following dispersion equation for the system under consideration:

$$\frac{\sum_{s=-\infty}^{\infty} [Q_s^{(2)} S_{0,s}^e (S_{0,s}^e)^* \operatorname{sh}^2 \Gamma_0^e h \operatorname{th} k_{y,s} y_0 + R_0^{e\,(2)} S_{0,s}^e (\tilde{S}_{0,s}^e)^* \operatorname{sh} \Gamma_0^e h \operatorname{ch} \Gamma_0^e h]}{\sum_{s=-\infty}^{\infty} [i\beta_s Q_s^{(1)} S_{0,s}^m (S_{0,s}^e)^* \operatorname{sh} \Gamma_0^e h \operatorname{ch} \Gamma_0^m h - R_0^{m\,(1)} T_{0,s}^m (S_{0,s}^e)^* \operatorname{sh} \Gamma_0^e h \operatorname{ch} \Gamma_0^m h]} =$$

$$= \frac{\sum_{s=-\infty}^{\infty} [-i\beta_s Q_s^{(1)} (S_{0,s}^m)^* S_{0,s}^e \operatorname{sh} \Gamma_0^e h \operatorname{ch} \Gamma_0^m h + R_0^{e\,(1)} (S_{0,s}^m)^* T_{0,s}^e \operatorname{sh} \Gamma_0^e h \operatorname{ch} \Gamma_0^m h]}{\sum_{s=-\infty}^{\infty} [(Q_s^{(2)})^* S_{0,s}^m (S_{0,s}^m)^* \operatorname{ch}^2 \Gamma_0^m h \operatorname{cth} k_{y,s} y_0 + (R_0^{m\,(2)})^* S_{0,s}^m (S_{0,s}^m)^* \operatorname{sh} \Gamma_0^m h \operatorname{ch} \Gamma_0^m h]}. \qquad (26)$$

It is seen from this equation that it contains two solutions, one of which describes the dispersion properties of a wave close to a TE wave (quasi-TE mode) while the other describes the dispersion properties of the quasi-TM mode. Great interest apparently attaches to the quasi-TM mode. In one particular case, which is not realized in practice this breaks up into two independent equations.

Further simplification can be attained by considering in regions I and II only the zeroth spatial harmonic. In this case we obtain the following second-order equation:

$$R_0^{e\,(2)} Q_0^{(2)} \tilde{S}_{0,0}^e S_{0,0}^m \operatorname{cth}^2 k_{y,0} y_0 + R_0^{m\,(2)} Q_0^{(2)} S_{0,0}^e S_{0,0}^m \operatorname{th} \Gamma_0^e h \operatorname{th} \Gamma_0^m h +$$
$$+ \operatorname{cth} k_{y,0} y_0 \{R_0^{e\,(2)} R_0^{m\,(2)} \tilde{S}_{0,0}^e S_{0,0}^m \operatorname{th} \Gamma_0^m h + [Q_0^{(2)} Q_0^{(2)} S_{0,0}^e S_{0,0}^m +$$
$$+ \beta_0^2 Q_0^{(1)} Q_0^{(1)} S_{0,0}^e S_{0,0}^m - i\beta_0 Q_0^{(1)} R_0^{m\,(1)} T_{0,0}^m S_{0,0}^e + i\beta_0 Q_0^{(1)} R_0^{e\,(1)} T_{0,0}^e S_{0,0}^m +$$
$$+ R_0^{m\,(1)} R_0^{e\,(1)} T_{0,0}^m T_{0,0}^e] \operatorname{th} \Gamma_0^e h\} = 0, \qquad (27)$$

which when solved with respect to $\coth k_{y,0} y_0$ yields a dispersion equation for symmetrical quasi-TE and quasi-TM modes.

The computations using these equations can be carried out in the following fashion: we fix the dimensions of the waveguide, the thickness and permittivity of the dielectric, and also the distance between diaphragms, we plot in accordance with (1) and (2) a family of curves characterizing the dependence of k_3 on Γ^e and Γ^m for different k. By choosing the necessary k_3 and determining from the resultant curves the values of Γ^e and Γ^m corresponding to the chosen k, we can determine the value of the height of the diaphragm h, at which the quasi-TE or quasi-TM mode will propagate with the necessary phase velocity.

REFERENCES

1. R.B.R. Shersby-Harvie, L.B. Mullet, W. Walkinshaw, S. Bell, G. Loach, Proc. I.E.E., 1957, 104B, 273.

2. N.A. Khizhnyak, Radiotekhnika i Elektronika, 1960, 5, 3, 413.
3. A.N. Didenko, E.S. Kovalenko, Atomnaya Energiya, 1961, 10, 1, 69.
4. L.A. Vaynshteyn, Electromagnetic Waves, Izd. Sovetskoye radio, 1957.

Submitted to the Editors, 21 January 1961

ON THE RITZ METHOD FOR HOLLOW SYSTEMS WITH ANISOTROPIC MEDIA

V.V. Nikol'skiy and V.G. Sukhov

Complete systems of eigenfunctions, corresponding to fields in unfilled waveguides and cavities, are used in the Ritz method as coordinate functions for waveguides and cavities with inhomogeneous anisotropic media (non-Hermitian permittivity tensors).

INTRODUCTION

The purpose of the present paper is to obtain for waveguides and cavities filled with an anisotropic medium the results of a successive application of the Ritz method. We have in mind here first of all the use of systems of coordinate functions, having the completeness property. It is natural to take systems of eigenfunctions, corresponding to the fields of empty (filled with homogeneous isotropic medium) cavities and waveguides. This method is particularly interesting because it should disclose a correspondence with the results of the direct eigenfunction expansion of the fields [1, 2]. It is shown below that under suitable matching and formulation of the problems (choice of type of expansions and form of the variational principle), the Ritz method leads to precisely the same results as the direct expansion of the fields. The equivalence of the two methods, established in this fashion, may be useful for various estimates. In addition to this fact, the clarity of the Ritz method as applied to these problems is worthy of attention. Because of this, for example for a waveguide, in addition to the already known results [2] presented here in clearer form, we have obtained new results, which are more convenient for computation.

1. THE RESONATOR PROBLEM

We shall consider a hollow electromagnetic resonator with inhomogeneous anisotropic medium (that is, containing in practice an anisotropic body, such as a magnetized ferrite). Starting with the variational equation

$$\omega^2 = \frac{\int_V \tilde{\vec{D}} \operatorname{rot} \mu^{-1} \operatorname{rot} \varepsilon^{-1} \vec{D}\, dv}{\int_V \tilde{\vec{D}} \vec{D}\, dv} \tag{1}$$

(Formula (8b) of [3]), we construct a Ritz scheme using the coordinate functions $\vec{E}_n$, corresponding to the free oscillations of the hollow cavity. Taking thus the expansions

$$\vec{D} = \varepsilon_0 \sum_n a_n \vec{E}_n, \quad \tilde{\vec{D}} = \varepsilon_0 \sum_n \tilde{a}_n \vec{E}_n^*, \tag{2}$$

we introduce them into the original formulation (1). The substitution leads to the following expression for ω^2 in terms of the bilinear forms:

$$\omega^2 = \frac{(\tilde{a}, \mathfrak{E}a)}{(\tilde{a}, a)}, \tag{3}$$

where $a = \{a_n\}$; $\tilde{a} = \{\tilde{a}_n\}$ are vectors made up of the coefficients of (2), with E_n orthonormalized as in [1]; the round brackets denote the following:

$$(\tilde{a}, a) = \sum_n \tilde{a}_n a_n, \quad (\tilde{a}, \mathfrak{E}a) = \sum_{k,n} \tilde{a}_k \mathfrak{E}_{kn} a_n$$

with

$$\mathfrak{E}_{kn} = \varepsilon_0 \int_V \vec{E}_k^* \operatorname{rot} \mu^{-1} \operatorname{rot} \varepsilon^{-1} \vec{E}_n \, dv.$$

Setting up the derivatives by the Ritz method

$$\frac{\partial}{\partial \tilde{a}_n} \{(\tilde{a}, \mathfrak{E}a) - \omega^2 (\tilde{a}, a)\}, \quad (n = 1, 2, 3, \ldots) \tag{4}$$

and equating them to zero, we obtain a system of equations equivalent to

$$\mathfrak{E}a = \omega^2 a. \tag{5}$$

This result exactly coincides with that in the "notes added in proof" in article [2] (Section 1). Precisely in the same way we obtain from a variational relationship containing the magnetic induction $\vec{B}$ given in the same reference,

$$\mathfrak{M}b = \omega^2 b, \tag{6}$$

where $b = \{b_n\}$ (b_n — the coefficients of expansion of $\vec{B}$ in the eigenfunctions $\vec{H}_n$) and

$$\mathfrak{M}_{kn} = \mu_0 \int_V \vec{H}_k^* \operatorname{rot} \varepsilon^{-1} \operatorname{rot} \mu^{-1} \vec{H}_n \, dv.$$

As regards the results of [1], they can be obtained from the variational principle in the form

$$\omega = j \frac{\int_V \left(\tilde{\vec{H}} \operatorname{rot} \vec{E} - \tilde{\vec{E}} \operatorname{rot} \vec{H}\right) dv}{\int_V \left(\tilde{\vec{H}} \mu \vec{H} + \tilde{\vec{E}} \varepsilon \vec{E}\right) dv} \tag{7}$$

(Formula (2) of [3]). To employ the Ritz method, we write, for example, the expansions

$$\begin{aligned} \varepsilon \vec{E} &= \varepsilon_0 \sum_n a_n \vec{E}_n, \\ \tilde{\vec{E}} &= \sum_n \tilde{a}_n \vec{E}_n^* + \sum_{n'} a_{n'} \vec{E}_{n'}^*; \\ \mu \vec{H} &= \mu_0 \sum_n b_n \vec{H}_n, \\ \tilde{\vec{H}} &= \sum_n \tilde{b}_n \vec{H}_n^* + \sum_{n'} \tilde{b}_{n'} \vec{H}_{n'}^*, \end{aligned} \tag{8}$$

where $\vec{E}_n$ and $\vec{H}_n$ are as before the vortical eigenfunctions of the cavity V, and $\vec{E}_{n'}$ and $\vec{H}_{n'}$ are the corresponding potential eigenfunctions. We further introduce, as was done before, the vectors $a = \{a_n\}$, $b = \{b_n\}$, $\tilde{a} = \{\tilde{a}_n\}, \tilde{a}' = \{\tilde{a}_{n'}\}$, $\tilde{b} = \{\tilde{b}_n\}$ and $\tilde{b}' = \{\tilde{b}_{n'}\}$. Inserting (8) into the variational relation (7), we reduce it to the form

$$\omega = \frac{(\tilde{a}, \Omega \mathfrak{I} a) + (\tilde{b}, \Omega M b)}{(\tilde{a}, a) + (\tilde{b}, b)}, \tag{9}$$

where $\Omega = [\omega_n]$ —diagonal matrix formed by the natural frequencies of the hollow resonator ω_n, and the elements of the matrices ϑ and M are the integrals

$$\vartheta_{kn} = \varepsilon_0 \int_V \vec{E}_k^* \varepsilon_0 \varepsilon^{-1} \vec{E}_n \, dv \text{ and } M_{kn} = \mu_0 \int_V \vec{H}_k^* \mu_0 \mu^{-1} \vec{H}_n \, dv.$$

We then obtain by the Ritz scheme

$$\begin{aligned} &\frac{\partial}{\partial \tilde{a}_n} \{(\tilde{a}, \Omega\vartheta a) + (\tilde{b}, \Omega M b) - \omega (\tilde{b}, b) - \omega (\tilde{a}, a)\} = 0 \\ &\qquad\qquad (n = 1, 2, 3, \ldots), \\ &\frac{\partial}{\partial \tilde{b}_n} \{(\tilde{a}, \Omega\vartheta a) + (\tilde{b}, \Omega M b) - \omega (\tilde{b}, b) - \omega (\tilde{a}, a)\} = 0 \end{aligned} \tag{10}$$

from which follow the matrix equations

$$\begin{aligned} \Omega\vartheta a &= \omega b, \\ \Omega M b &= \omega a. \end{aligned} \tag{11}$$

Excluding a and b in succession, we obtain finally

$$\begin{aligned} \mathfrak{E} a &= \omega^2 a, \\ \mathfrak{M} b &= \omega^2 b, \end{aligned} \tag{12}$$

where

$$\begin{aligned} \mathfrak{E} &= \Omega M \Omega \vartheta; \\ \mathfrak{M} &= \Omega \vartheta \Omega M. \end{aligned}$$

Equations (11) and (12) coincide with those obtained in [1] [formulas (9'') and (10')]. We have merely here slightly simplified the notation by writing the integrals of the matrix elements ϑ_{kn} and M_{kn} for the total volume of the resonator: the perturbation region is not singled out.

Inverse relations of the form

$$\begin{aligned} \mathfrak{E}^{-1} a &= \omega^{-2} a, \\ \mathfrak{M}^{-1} b &= \omega^{-2} b \end{aligned} \tag{13}$$

are also derived from (7) by using the voltage expansions only. We note that comparison with Equations (18) of [1] lead to certain simplifications, when the properties of the divergences $\operatorname{div} \mu^{-1} \operatorname{rot} \vec{E}$ and $\operatorname{div} \varepsilon^{-1} \operatorname{rot} \vec{H}$ are taken into account.

2. THE WAVEGUIDE PROBLEM (EXPANSION OF FIELD INTENSITIES)

By way of a convenient formulation of the variational principle for a waveguide with anisotropic longitudinally homogeneous filling, we write down the expression

$$\Gamma = \frac{\omega \int_{S_\perp} \left(\tilde{\vec{E}} \varepsilon \vec{E} + \tilde{\vec{H}} \mu \vec{H}\right) ds - j \int_{S_\perp} \left(\tilde{\vec{H}} \operatorname{rot}_\perp \vec{E} - \tilde{\vec{E}} \operatorname{rot}_\perp \vec{H}\right) ds}{\int_{S_\perp} \{[\vec{E}, \tilde{\vec{H}}] + [\tilde{\vec{E}}, \vec{H}]\} \vec{ds}} \tag{14}$$

(Formula (10) of [3]).

Since it is not our purpose here to obtain results which are already known, we start from the field intensities. Using henceforth the notation and normalization * of [2], we take

* We note that the change in normalization (namely the normalization to the unit of the transverse functions) on change in designation for the transverse fields (for example, the index I for H fields and I' for E fields) introduces certain simplifications in the notation.

$$\vec{E} = \sum_m (a_{zm}\vec{E}_{zm} + a_{\perp m}\vec{E}_{\perp m}) + \sum_n a_{\perp' n}\vec{E}_{\perp' n},$$
$$\tilde{\vec{E}} = \sum_m (\tilde{a}_{zm}\vec{E}^*_{zm} + \tilde{a}_{\perp m}\vec{E}^*_{\perp m}) + \sum_n \tilde{a}_{\perp' n}\vec{E}^*_{\perp' n}; \quad (15')$$
$$\vec{H} = \sum_n (b_{zn}\vec{H}_{zn} + b_{\perp n}\vec{H}_{\perp n}) + \sum_m b_{\perp' m}\vec{H}_{\perp' m},$$
$$\tilde{\vec{H}} = \sum_n (\tilde{b}_{zn}\vec{H}^*_{zn} + \tilde{b}_{\perp n}\vec{H}^*_{\perp n}) + \sum_m \tilde{b}_{\perp' m}\vec{H}^*_{\perp' m}. \quad (15'')$$

In accordance with expansions (15), we introduce the vectors a, a, b, and b, each of which is made up of all the coefficients of one of the lines (15), so that, for example, a is the unification of the vectors $a_z = \{a_{zm}\}$, $a_1 = \{a_{1m}\}$, and $a_{1'} = \{a_{1'n}\}$. Substitution of these expansions in (14) yields

$$\Gamma = \frac{(\tilde{a}, \mathcal{A}a) + (\tilde{b}, M b) - (\tilde{b}, X^2 a) - (\tilde{a}, X^2 b)}{(\tilde{a}, \mathring{\Gamma} b) + (\tilde{b}, \mathring{\Gamma} a)}. \quad (16)$$

The parentheses denote here bilinear forms, as previously. As regards the operators in (16), they are conveniently broken up into cells in accordance with the scheme

$$\begin{Vmatrix} zz & z\perp & z\perp' \\ \perp z & \perp\perp & \perp\perp' \\ \perp' z & \perp'\perp & \perp'\perp' \end{Vmatrix}.$$

Then

$$\mathcal{A} = \begin{Vmatrix} \mathcal{A}_{zz} & \mathcal{A}_{z\perp} & \mathcal{A}_{z\perp'} \\ \mathcal{A}_{\perp z} & \mathcal{A}_{\perp\perp} & \mathcal{A}_{\perp\perp'} \\ \mathcal{A}_{\perp' z} & \mathcal{A}_{\perp'\perp} & \mathcal{A}_{\perp'\perp'} \end{Vmatrix}, \qquad M = \begin{Vmatrix} M_{zz} & M_{z\perp} & M_{z\perp'} \\ M_{\perp z} & M_{\perp\perp} & M_{\perp\perp'} \\ M_{\perp' z} & M_{\perp'\perp} & M_{\perp'\perp'} \end{Vmatrix},$$

where

$$\mathcal{A}_{\alpha\beta} = \left\| \int_{S_\perp} \vec{E}^*_{\alpha k} \varepsilon \vec{E}_{\beta n}\, ds \right\|; \quad M_{\alpha\beta} = \left\| \int_{S_\perp} \vec{H}^*_{\alpha k} \mu \vec{H}_{\beta n}\, ds \right\| \quad (\alpha, \beta = z, \perp, \perp')$$

and

$$X^2 = \frac{1}{k_0^2} \begin{Vmatrix} 0 & 0 & [\chi_n^2] \\ 0 & 0 & 0 \\ [\chi_n^2] & 0 & 0 \end{Vmatrix}; \quad \mathring{\Gamma} = \frac{1}{k_0^2} \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & [\Gamma_n] \\ 0 & [\Gamma_n] & 0 \end{Vmatrix};$$

$[\chi_n^2]$ and $[\Gamma_n]$ are diagonal matrices formed by the transverse and longitudinal wave numbers of different waves of a empty waveguide (the notation is from [2]!). Let M be the numerator and N the denominator of the right half of (16). By the Ritz method

$$\frac{\partial}{\partial \tilde{a}_{an}}(M - \Gamma N) = 0$$
$$\frac{\partial}{\partial \tilde{b}_{an}}(M - \Gamma N) = 0 \qquad (a = z, \perp, \perp';\ n = 1, 2, 3, \ldots), \quad (17)$$

which yields two systems of equations, which have the following matrix form

$$\mathcal{A}a - X^2 b = \Gamma \mathring{\Gamma} b,$$
$$Mb - X^2 a = \Gamma \mathring{\Gamma} a. \quad (18)$$

Eliminating b and a in succession, we obtain

$$a = \mathcal{A}^{-1}(\Gamma\mathring{\Gamma} + X^2) M^{-1} (\Gamma\mathring{\Gamma} + X^2) a,$$
$$b = M^{-1}(\Gamma\mathring{\Gamma} + X^2) \mathcal{A}^{-1} (\Gamma\mathring{\Gamma} + X^2) b. \quad (19)$$

An equation in the unknown propagation constant Γ (the so called "dispersion equation") is obtained by letting the corresponding determinant vanish:

$$\begin{aligned} &\mathrm{Det}\,|\,\mathcal{Э} - (\Gamma\mathring{\Gamma} + X^2)\,M^{-1}\,(\Gamma\mathring{\Gamma} + X^2)\,| = 0,\\ &\mathrm{Det}\,|\,M - (\Gamma\mathring{\Gamma} + X^2)\,\mathcal{Э}^{-1}\,(\Gamma\mathring{\Gamma} + X^2)\,| = 0. \end{aligned} \tag{20}$$

A particularly lucid formulation is obtained by combining the vectors a and b into a new vector $c = \{a, b\}$ and accordingly introducing the "total energy" operator w such that $Wc = \{\mathcal{Э}a, Mb\}$ *. Then the unification of equations (18)

$$\begin{Vmatrix} \mathcal{Э} & 0 \\ 0 & M \end{Vmatrix} \begin{pmatrix} a \\ b \end{pmatrix} - \begin{Vmatrix} 0 & X^2 \\ X^2 & 0 \end{Vmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \Gamma \begin{Vmatrix} 0 & \hat{\Gamma} \\ \mathring{\Gamma} & 0 \end{Vmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{18a}$$

has the natural form

$$(W - \check{X}^2)\,c = \Gamma\check{\Gamma}c \tag{21}$$

[The meaning of the operators $\check{X}^2$ and $\check{\Gamma}$ is clear from (18a)], and the dispersion equation assumes the form

$$\mathrm{Det}\,|\,W - \check{X}^2 - \Gamma\check{\Gamma}\,| = 0, \tag{22}$$

which is very similar to the elementary dispersion relation

$$k_0^2 - \chi^2 - \Gamma^2 = 0.$$

It is easy to see that when $\varepsilon = I\varepsilon_0$ and $\mu = I\mu_0$ (I —(I is a unit tensor) Equations (22) breaks up into a series of such relationships. Equation (22) in general shows that the quantity $1/\Gamma$ runs runs through the eigenvalues of the matrix

$$(W - \check{X}^2)^{-1}\,\check{\Gamma}.$$

By way of an example of the appreciable simplification of the final formulation (as a result of which its computational value is greatly enhanced) let us consider the case when the material tensors have the form

$$\varepsilon = \begin{pmatrix} \varepsilon_1 & -j\varepsilon_2 & 0 \\ j\varepsilon_2 & \varepsilon_1 & 0 \\ 0 & 0 & \varepsilon_z \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 & -j\mu_2 & 0 \\ j\mu_2 & \mu_1 & 0 \\ 0 & 0 & \mu_z \end{pmatrix}$$

(z is the longitudinal axis of the waveguide!). In this case the cells $\perp z$, $\perp' z$, $z \perp$ и $z \perp'$ and z_1' in the matrices $\mathcal{Э}$ and M vanish, so that we obtain from the first cell rows of (18)

$$\mathcal{Э}_{zz}a_z - \frac{[\chi_n^2]}{k_0^2}\,b_{\perp'} = 0, \quad M_{zz}b_z - \frac{[\chi_n^2]}{k_0^2}\,a_{\perp'} = 0, \tag{23}$$

and elimination of a_z and b_z from the second and third rows yields

$$\mathcal{Э}'a' = \Gamma\mathring{\Gamma}b', \quad M'b' = \Gamma\mathring{\Gamma}a'. \tag{24}$$

In this notation the primes of the vectors a and b denote that their z components are discarded, and the matrices $\mathcal{Э}'$ and M' are specified in the following fashion:

---

* The same operations can be carried out also for a resonator on the basis of (11).

$$\mathcal{Э}' = \left\| \begin{matrix} \mathcal{Э}_{\perp\perp} & \mathcal{Э}_{\perp\perp'} \\ \mathcal{Э}_{\perp'\perp} & \mathcal{Э}_{\perp'\perp'} - \frac{1}{k_0^4} [\chi_n^2] M_{zz}^{-1} [\chi_n^2] \end{matrix} \right\|,$$

$$M' = \left\| \begin{matrix} M_{\perp\perp} & M_{\perp\perp'} \\ M_{\perp'\perp} & M_{\perp'\perp'} - \frac{1}{k_0^4} [\chi_n^2] \mathcal{Э}_{zz}^{-1} [\chi_n^2] \end{matrix} \right\|.$$

With this, as can be readily shown

$$\frac{1}{k_0^4} [\chi_n^2] M_{zz}^{-1} [\chi_n^2] = \left\| \mu_0 \int_{S_\perp} \vec{H}_{zk}^* \mu_0 \mu_z^{-1} \vec{H}_{zn} \, ds \right\|,$$

$$\frac{1}{k_0^4} [\chi_n^2] \mathcal{Э}_{zz}^{-1} [\chi_n^2] = \left\| \varepsilon_0 \int_{S_\perp} \vec{E}_{zk}^* \varepsilon_0 \varepsilon_z^{-1} \vec{E}_{zn} \, ds \right\|.$$

From (24) we eliminated b' and a'. The equations resulting from this

$$\mathring{\Gamma}^{-1} M' \mathring{\Gamma}^{-1} \mathcal{Э}' \, a' = \Gamma^2 a', \qquad (25)$$

$$\mathring{\Gamma}^{-1} \mathcal{Э}' \mathring{\Gamma}^{-1} M' b' = \Gamma^2 b',$$

reduce the definition of Γ to the determination of the eigenvalues of relatively simple matrices.

3. THE WAVEGUIDE PROBLEM (EXPANSION OF INDUCTIONS)

The formulation obtained above for the waveguide coincides with the results of [2]. However, the variational relationship (14) yields the previous result if we take the expansions

$$\varepsilon \vec{E} = \varepsilon_0 \sum_m (a_{zm} \vec{E}_{zm} + a_{\perp m} \vec{E}_{\perp m}) + \varepsilon_0 \sum_n a_{\perp' n} \vec{E}_{\perp' n},$$

$$\tilde{\vec{E}} = \sum_m (\tilde{a}_{zm} \vec{E}_{zm}^* + \tilde{a}_{\perp m} \vec{E}_{\perp m}^*) + \sum_n \tilde{a}_{\perp' n} \vec{E}_{\perp' n}^*; \qquad (26')$$

$$\mu \vec{H} = \mu_0 \sum_n (b_{zn} \vec{H}_{zn} + b_{\perp n} \vec{H}_{\perp n}) + \mu_0 \sum_m b_{\perp' m} \vec{H}_{\perp' m},$$

$$\tilde{\vec{H}} = \sum_n (\tilde{b}_{zn} \vec{H}_{zn}^* + \tilde{b}_{\perp n} \vec{H}_{\perp n}^*) + \sum_m \tilde{b}_{\perp' m} \vec{H}_{\perp' m}^*. \qquad (26'')$$

We then obtain from (14) the following expression for Γ in terms of bilinear forms:

$$\Gamma = \frac{(\tilde{a}, Na) + (\tilde{b}, Nb) - (\tilde{b}, X^2 \mathcal{Э} a) - (\tilde{a}, X^2 M b)}{(\tilde{b}, \mathring{\Gamma} \mathcal{Э} a) + (\tilde{a}, \mathring{\Gamma} M b)}, \qquad (27)$$

where unlike (16) the cells in the matrices $\mathcal{Э}$ and M have the form

$$\mathcal{Э}_{\alpha\beta} = \left\| \varepsilon_0 \int_{S_\perp} \vec{E}_{\alpha k}^* \varepsilon^{-1} \varepsilon_0 \vec{E}_{\beta n} \, ds \right\|, \quad M_{\alpha\beta} = \left\| \mu_0 \int_{S_\perp} \vec{H}_{\alpha k}^* \mu^{-1} \mu_0 \vec{H}_{\beta n} \, ds \right\| (\alpha, \beta = z, \perp, \perp');$$

the matrices X^2 and $\mathring{\Gamma}$ are

$$X^2 = \frac{1}{k_0^2} \left\| \begin{matrix} 0 & 0 & [\chi_n^2] \\ 0 & 0 & 0 \\ [k_0^2] & 0 & 0 \end{matrix} \right\|, \quad \mathring{\Gamma} = \frac{1}{k_0^2} \left\| \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & [\Gamma_n] \\ 0 & k_0^2 [\Gamma_n^{-1}] & 0 \end{matrix} \right\|,$$

and in addition, we have also the normalization matrix N:

$$N = \frac{1}{k_0^2} \left\| \begin{matrix} [\chi_n^2] & 0 & 0 \\ 0 & [\Gamma_n^2] & 0 \\ 0 & 0 & [k_0^2] \end{matrix} \right\|.$$

From (27) we obtain subject to conditions such as (17)

$$Na = (X^2 + \Gamma\mathring{\Gamma})\, Mb, \quad Nb = (X^2 + \Gamma\mathring{\Gamma})\, \partial a, \tag{28}$$

and further, by eliminating b and a we get

$$\begin{aligned} N^{-1}(X^2 + \Gamma\mathring{\Gamma})\, MN^{-1}(X^2 + \Gamma\mathring{\Gamma})\, \partial a = a, \\ N^{-1}(X^2 + \Gamma\mathring{\Gamma})\, \partial N^{-1}(X^2 + \Gamma\mathring{\Gamma})\, Mb = b. \end{aligned} \tag{29}$$

On the basis of (29) we can easily write a dispersion equations similar to (20). Unlike (19), however, we can eliminate from (29) the components a_z and b_z. Indeed, the first and second cell rows of the first equation of (28) have the form

$$\begin{aligned} a_z &= M_{\perp' z} b_z + M_{\perp' \perp} b_\perp + M_{\perp' \perp'} b_{\perp'}, \\ \frac{[\Gamma_n]}{\Gamma} a_\perp &= M_{\perp' z} b_z + M_{\perp' \perp} b_\perp + M_{\perp' \perp'} b_\perp, \end{aligned}$$

that is, they imply that

$$a_{zn} = \frac{\Gamma_n}{\Gamma} a_{\perp n}. \tag{30}$$

In exactly the same manner we obtain from the first two cell rows of the second equation of (28)

$$b_{zn} = \frac{\Gamma_n}{\Gamma} b_{\perp n}. \tag{31}$$

Leaving in each equation of (28) the second and third cell rows, we obtain matrix equations with respect to the vectors a' and b', containing no z components:

$$\begin{aligned} \partial' a' &= b', \\ M' b' &= a'. \end{aligned} \tag{32}$$

Here

$$\partial' = \begin{Vmatrix} [\Gamma_n^{-1}]\, \partial_{\perp' z} [\Gamma_n] + \Gamma [\Gamma_n^{-1}]\, \partial_{\perp' \perp} & \Gamma [\Gamma_n^{-1}]\, \partial_{\perp' \perp'} \\ (\partial_{zz} + \Gamma [\Gamma_n^{-1}]\, \partial_{\perp z}) \frac{[\Gamma_n]}{\Gamma} + \partial_{z\perp} + \Gamma [\Gamma_n^{-1}]\, \partial_{\perp \perp} & \partial_{z\perp'} + \Gamma [\Gamma_n^{-1}]\, \partial_{\perp \perp'} \end{Vmatrix};$$

$$M' = \begin{Vmatrix} [\Gamma_n^{-1}]\, M_{\perp' z} [\Gamma_n] + \Gamma [\Gamma_n^{-1}]\, M_{\perp \perp'} & \Gamma [\Gamma_n^{-1}]\, M_{\perp' \perp'} \\ (M_{zz} + \Gamma [\Gamma_n^{-1}] M_{\perp z}) \frac{[\Gamma_n]}{\Gamma} + M_{z\perp} + \Gamma [\Gamma_n^{-1}]\, M_{\perp \perp} & M_{z\perp'} + \Gamma [\Gamma_n^{-1}]\, M_{\perp \perp'} \end{Vmatrix}.$$

The dispersion equation that follows from (32) has the form

$$\mathrm{Det}\,|\,\partial' M' - I\,| = 0 \quad \text{or} \quad \mathrm{Det}\,|\,M' \partial' - I\,| = 0. \tag{33}$$

A direct comparison shows that all these results coincide with those obtained in [2]. In particular, formulas (30) and (31) correspond to (36) and (37) of [2].

CONCLUSION

We have shown above that a successive application of the Ritz method to the corresponding variational formulation for a resonator and for a waveguide leads to the same results as does the direct eigenfunction expansion of the fields. However, the improvement in formalism, obtained in the present paper, has not only yielded the earlier results in a more lucid form, but made it possible to obtain for the case considered in Section 2 an expression which is more convient from the computational point of view.

REFERENCES

1. V. V. Nikol'skiy, Investigation of Hollow Systems with Anisotropic Regions by the Eigenfunction Method, Part I. Resonator. Radiotekhnika i Elektronika, 1960, 5, 11, 1802.
2. V. V. Nikol'skiy, Investigation of Hollow Systems with Anisotropic Regions by the Eigenfunction Method, Part III, Waveguide. Radiotekhnika i Elektronika, 6, 1, 74.
3. V. V. Nikol'skiy, Variational Principle for Hollow Systems with Anisotropic Medium, Radiotekhnika i Elektronika, 1960, 6, 9, 1583.

Submitted to the Editors 10 January 1961.

———

ADVISABLE FORM OF THE EXPRESSION FOR THE COUPLING COEFFICIENT IN THE INTERACTION BETWEEN AN ELECTRON BEAM AND SPATIAL HARMONICS OF A SLOW-WAVE SYSTEM

L.N. Loshakov

Arguments are advanced, showing the advisability of modifying the previously introduced expression for the coupling coefficient in the interaction between an electron beam and an arbitrary spatial harmonic in a system with periodic structure.

In [1,2], devoted to the derivation of the characteristic equation for traveling-wave and backward-wave tubes with the aid of the Lorentz lemma, the sought propagation constants in the system with an electron beam were defined by the expression

$$\gamma_m = \mp j\beta_{00}(1 \pm \delta) \mp j\frac{2\pi}{h} m = \mp j\beta_{0m}\left(1 \pm \frac{\beta_{00}}{\beta_{0m}}\delta\right). \quad (1)$$

In this expression β_{00} and β_{0m} denote the phase constant of the zeroth and or the m-th spatial harmonics in the slow-wave system having a structure period h in the absence of the beam; δ is the additional term which accounts for the relative change in the phase constant of the zeroth harmonic, due to the presence of the electron beam. The upper sign pertains to the case of interaction with the forward spatial harmonic, and the lower one to the case of interaction with the backward spatial harmonic.

The characteristic equation for the additional term δ in the interaction with the n-th spatial harmonic of the complex wave existing in the system was derived in the form

$$\delta\left[\left(\delta + \frac{1-\eta}{\chi}\right)^2 - \Gamma M\right] \pm M\frac{\omega\varepsilon}{4\beta_{00}P_0}\int_{s_0} |E_{0nz}|^2 ds = 0. \quad (2)$$

In Equation (2) $\chi = \frac{\beta_{00}}{|\beta_{0n}|}$; η is the ratio of the phase velocity of the interacting (working) harmonic to the mean velocity of the electron; $M = q/\beta_{00}^2$ is the current parameter; q the square of plasma wave number of an infinitely broad beam; Γ the dispersion coefficient of the space-charge field; P_0 and $|E_{0nz}|$ are the time-averaged total power flux of the complex wave and the amplitude of the electric field of the working harmonic in the direction of possible motion of electrons, existing in the absence of the electron beam; ω is the angular frequency; ϵ is the dielectric constant of the medium, and s_0 is the cross-sectional area of the electron beam. The rational practical (MKS) system of units was used in the calculations.

The dimensionless quantity

$$K_{cn} = \pm\frac{\omega\varepsilon}{2\beta_{00}P_0}\int_{s_0} |E_{0nz}|^2 ds \quad (3)$$

was called the coupling coefficient in the interaction with the n-th forward or backward (lower-sign) spatial harmonic. The proposed expression for K_{cn} cannot be regarded as fully successful by virtue of the fact that according to formula (1) the relative perturbation of the phase constant of the working (n-th) spatial harmonic, characterizing the efficiency of interaction with it, is accounted for by the additional term $\frac{\beta_{00}}{|\beta_{0n}|}\delta^*$. Consequently, expression

* For the backward working harmonic we have $\beta_{0n} < 0$.

(3) represents in the general case (with the exception of N = 0) an intermediate parameter, which determines the change in this harmonic due to the interaction with the electron beam less completely than if K_{cn} were to be introduced as applied to the characteristic equation for the additional term of the working harmonic.

It is easily seen that the characteristic equation for the additional term $\delta' = \frac{\beta_{00}}{|\beta_{0n}|}\delta$ can be obtained by multiplying Equation (2) by $\left(\frac{\beta_{00}}{|\beta_{0n}|}\right)^3$. Then the new characteristic equation will have the following form in the previous notation.

$$\delta'[(\delta' + 1 - \eta)^2 - \Gamma M] + \frac{1}{2}MK_{cn} = 0, \tag{4}$$

where

$$\eta = \frac{\omega}{u_0|\beta_{0n}|}; \quad M = \frac{q}{\beta_{0n}^2} = \frac{ei_0}{meu_0^3\beta_{0n}^2}; \tag{5}$$

$$K_{cn} = \pm\frac{\omega\varepsilon}{2|\beta_{0n}|P_0}\int_{s_n}|E_{0nz}|^2\,ds. \tag{6}$$

In formulas (5), e and m are the charge and mass of the electron; i_0 and u_0 are respectively the DC components of the current density and of the electron velocity.

The coupling coefficient, defined by formula (6), describes more fully the perturbed propagation constant of the working spatial harmonic, and consequently the efficiency of interaction, that is, the quality of the slow-wave system. This expression is invariant with respect to the numbering of the harmonics, corresponding to the actual arbitrariness in their numbering. If the calculation is modified as indicated, the propagation constant of the working harmonic in the system with an electron beam is given by the formula

$$\gamma_n = -j|\beta_{0n}|(1 + \delta'). \tag{7}$$

The propagation constants of all the spatial harmonics of a complex wave containing one working harmonic are subjected to the same increment $(-j|\beta_{0n}|\delta')$; therefore the following equation is valid for the propagation constants of the nonoperating harmonics of this complex wave:

$$\gamma_m = \mp j\beta_{0m}\left(1 \pm \frac{|\beta_{0n}|}{\beta_{0m}}\delta'\right). \tag{8}$$

Formula (8) shows that the relative perturbation introduced by the beam is different for differently numbered harmonics.

The need for having a coupling coefficient that is invariant under renumbering of the harmonics, so that it can be used as an independent parameter to estimate the efficiency of the slow-wave systems, was pointed out to the author by L. P. Lisovskiy.

In conclusion, it is of practical interest to note that the parameters η, M, and K_{cn} of Equation (4), defined by formulas (5) and (6), can be related to the velocity detuning parameter b, the amplification C, and the coupling resistance R_{cn}, used in Pierce's theory, with the aid of the equations

$$\eta = \frac{1}{1 + bC}, \tag{9}$$

$$K_{cn} = \omega\varepsilon|\beta_{0n}|s_0R_{cn}, \tag{10}$$

$$\frac{1}{2\eta}MK_{cn} = C^3, \tag{11}$$

and near synchronism we have $\eta \simeq 1$. In the derivation of (10) it was assumed that the expression for the coupling resistance includes the mean-squared value of the electric field of the interacting spatial harmonic, taken over the cross section of the beam.

The formulas (9) - (11) given above can be useful for a relative estimate of the results of different computation methods. In particular, K_{cn} characterizes apparently more accurately than R_{cn} the coupling between the electron beam and the slow-wave system, and takes the geometry of the interaction space into account. The value of this parameter is consequently unchanged if the systems are connected in parallel. By virtue of this, the dimensionless coupling coefficient can be regarded as more convenient for estimating the quality of slow-wave systems than the coupling resistence.

REFERENCES

1. L. N. Loshakov, The Use of the Lorentz Lemma for the Determination of the Propagation Constants in the Interaction Between an Electron Beam and Spatial Harmonics. Radiotekhnika i Elektronika, 1957, 2, 4, 461.
2. L. N. Loshakov, Considerations of the Space-Charge Field in the Calculation of the Propagation Constants in a Slow-Wave System in the Presence of an Electron Beam, Using the Lorentz Lemma. Radiotekhnika i Elektronika, 1959, 4, 4, 688.

Submitted to the Editors 24 March 1961

INFLUENCE OF PULSATIONS OF THE CROSS SECTION OF THE ELECTRON BEAM ON THE NOISE FACTOR OF A TRAVELING-WAVE TUBE

V. B. Glasko, A. A. Zyuzin-Zinchenko, V. M. Lopukhin

The effect of pulsations in the cross section of the electron beam in an electron gun on the noise factor of a traveling-wave tube is investigated. Plots are given showing the dependence of the noise factor on the parameters Δ and β, which characterize the pulsation of the beam cross section. The noise factor is plotted as a function of the length of the electron gun for different values of Δ and β. It has been established that the pulsations of the electron beam lead as a rule to an increase in the noise factor of the traveling-wave tube.

The question of calculating the noise factor of a TWT, which determines its sensitivity, has been considered in many articles, of which the most interesting are those devoted to a reduction of this factor. Articles [1,2,3] review this work up to 1958 and contain a detailed bibliography. Analogous questions were considered in the books [4,5].

Berghammer and Bloom [6] calculated the noise factor of a TWT taking into account the electron distribution function. By integrating the Boltzmann equation they have shown that the noise parameters S and π are invariant in electron beams. It is shown in the same paper that the multivelocity character of the electron beam leads to a reduction in the minimum attainable noise factor of a TWT.

References [7,8] consider the correlation of the current and velocity fluctuations on the cathode.

Robinson [7] has shown that the current and velocity fluctuations are uncorrelated in temperature-limited emission. On going to a space-charge limited current, the current fluctuations decrease; in addition, a correlation arises between the current and velocity fluctuations.

Tien and Moshman [8] simulated the random process near the potential minimum with the aid of the Monte Carlo method and arrived, under certain physical assumptions, at the conclusion that there is no correlation between the current and velocity fluctuations in the case of a space-charge limited current.

The discrepancy between the deductions of [7] and [8] is due to the different physical picture of the processes near the virtual cathode, assumed in these papers.

Gertsenshteyn [9] estimated the volume noise which should exist if the electron beam is regarded as a statistical system. He has shown that the volume noise can predominate over the noise due to shock processes on the cathode in the millimeter band.

Several papers [10,11] have been devoted to an experimental study of noise processes in the electron beam and to the noise factor of a TWT. MacIntosh [10] used grids and diaphragms moving along the electron beam to investigate experimentally the current-distribution noise.

Saito Shigebumi [11] described a method of experimentally measuring the real part of the correlation coefficient of the current and velocity fluctuations in the electron beam. Two resonators spaced $\lambda p/4$ were displaced along the beam (λp is the plasma wavelength of the electron beam). Microwave noise was applied to a twin-T bridge at the input of a modulation radiometer. Such a circuit made it possible to measure separately the noise in the fast and slow electron waves.

Many papers contain descriptions of new methods of reducing the noise factors of TWT and BWT in the amplification mode. Currie and Forster [12,13] investigated experimentally backward-wave tube amplifiers with low noise. They used in these tubes tubular electron beams with special distribution of the field along the beam in the guns. These guns were characterized by the absence of a potential minimum near the cathode, and also by a rather slow potential buildup in the region close to the cathode. In this region (where the beam has a multivelocity character with low mean value of the velocity) the noise in the electron beam is indeed suppressed. The minimum experimental value of the noise factor of BWT amplifiers with electron guns of this type was 2.9 db.

In [14] is described a TWT for the 10 cm band with a noise factor 3.5 db. The electron beam had a complex cross section, the field varied with the coordinate rather slowly, and the field distribution was the same as in the Currie and Forester guns.

Shaw and his co-workers [15] also pointed out the possibility of suppressing the noise in an electron beam by using a drift space with low potential. In an experimental tube for the 3 cm band with an eight-electrode gun and an L cathode, the minimum noise factor obtained was 6.7 db. In the case of an oxide cathode a reduction is expected in the noise factor.

Mueller and Currie [16] have shown that if the potential builds up fairly smoothly in the gun, noise factors close to unity can be obtained. The noise factor decreases with increasing frequency.

Gertsenshteyn [17] considered the space beats of noise waves in coupled slow-wave systems. He showed that the beats give rise to points where both the noise current and the noise voltage vanish. A construction of this type, placed between the electron gun and the input to the traveling-wave tube, can in principle greatly reduce the noise factor of the TWT.

Several papers have been devoted to a study of parametric TWT tubes with transverse field [18 — 21]. The minimum noise factor attained in such tubes was $F_{min} = 2.3$ db. A review of work done in this field is given in the article [22].

Summarizing the brief review of work on the reduction of the noise factor of TWT and BWT for centimeter wavelengths, it must be pointed out that the greatest success was attained in guns with rather slowly varying potential, and also in parametric TWT with transverse field. We note that in all the papers referred to above the electron beam is assumed to have a constant cross section in the gun.

The present article is devoted to a somewhat different group of problems. The point is that under practical conditions of shaping the electron beam, its cross section frequently pulsates and varies with the coordinates. The reasons for the pulsations may be disturbances to the axial symmetry of the gun, inhomogeneity of the magnetic field in the gun, and the radial velocity-components of the electrons emitted from the cathode.

The purpose of the article is a study of the influence of pulsations in the cross section of the electron beam in the gun on the noise factor of a TWT with a two-anode gun. In such a gun there is a potential minimum near the cathode, and as the distance from the cathode increases, the potential increases rather smoothly, thereby ensuring a minimum noise factor for a TWT with cylindrical beam, on the order of seven units. The variation of the potential is such a gun differs from that in multielectrode guns [12,13,14], with the aid of which lower noise factors were obtained in TWT and BWT. With such a formulation of the problem, the character of the correlation between the noise and velocity fluctuations in a certain initial gun cross section is not significant. It is therefore advantageous to make the simplest assumption, namely that there is no correlation between the fluctuations in the initial cross section of the gun. Formally, the correlation can then be taken into account following [27].

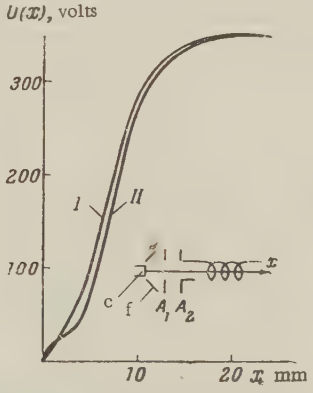


Figure 1. Diagram of TWT gun; static potential in the gun U(x) as a function of the coordinate x in modes I and II.

Let us consider a three-electrode TWT gun with cylindrical cathode C, having a focusing electrode F, and also anodes A_1 and A_2 (see Figure 1). We introduce a cylindrical system of coordinates x, r, θ, in which the x-axis coincides with the axis of the beam, and the plane $x = 0$ coincides with the virtual cathode (the coordinate x is measured in millimeters everywhere). The gun is in a longitudinal constant focusing magnetic field H_0. Theory and experiments show that in this case the electron beam is pulsating, and its cross section varies quasi-periodically with the coordinate x. The parameters characterizing the pulsation of the beam depend on the electrostatic field E_0 in the gun, on the magnetic field H_0, and on the ratio of the velocities v_x and v_r on the cathode. The form of the cross section of the electron beam can be approximately determined theoretically by calculating the trajectories of the electrons on the periphery of the beam in the fields H_0 and $E_0 = -\text{grad}\, U(x, r, \theta)$ where $U(x, r, \theta)$ is the static

potential in the gun, determined without taking the space charge into account. Integrating the equations of motion of the peripheral electron in the fields H_0 and E_0 for one of the working modes of the TWT with three-electrode gun, we can verify that the electron trajectory is given approximately by

$$b = b_0 [1 + \Delta \sin \beta k(x) x], \tag{1}$$

where b_0 is the radius of unperturbed beam; Δ and β are constants depending on H_0, E_0, and the initial conditions imposed on the electron velocities at the section x = 0; k(x) is a function of the form

$$k(x) = 820 (x + 6)^{-3} + 0.4. \tag{2}$$

We shall assume below that expression (1), with k(x) given by (2), describes the variation of the cross section of the electron beam. The parameter Δ characterizes the depth of modulation of the beam cross section, $\beta k(x)$ is the modulation wave number, so that $\Lambda(x) = 2\pi/\beta k(x)$ is the quasi-period of the modulation. We shall also assign to the parameters Δ and β different possible character of periodicity of the beam as a function of the initial conditions in the beam in the potential minimum at x = 0.

The purpose of the present paper is thus to establish the dependence of the noise factor F of a TWT on the parameters Δ_i and β_i (i = 1, 2, ..., 22) for two electric modes of gun operation, under which the distribution of the static potential in the gun is given by curves I and II of Fig. 1. The variants i = 1 — 16 correspond to mode I, while variants i = 17 — 22 correspond to mode II. In mode I the current is I_{01} whereas in mode II it is I_{02}. A potential distribution close to the curve I or II of Fig. 1 can be obtained, for example, by choosing proper voltages for the electrodes of the three-electrode gun (see, for example, reference [23]).

To calculate the noise factor of the TWT we assume that there are uncorrelated fluctuations of the current I and of the velocity v on the virtual cathode; these fluctuations then propagate in the electrostatic field of the gun, and the cross section of the beam varies with the coordinate x.

Let us assume that the potential U(x) varies with the coordinate x sufficiently slowly, so that no higher modes are excited in the beam.

We assume also that near the cathode, at distances $x < x_1$, where $x_1 = 0.2\, b_0$ (b_0 is the radius of the unperturbed beam in millimeters), the potential U(x) assumes the form obtained by Langmuir for the potential near the virtual cathode [24].

We note that the exact form of the field U(x) for $x < 0.2 b_0$ is not important for the deductions of the present paper, since its purpose is to establish the dependence of the noise factor on the pulsations in the cross section of the electron beam at a fixed distribution of the potential U(x).

Approximating an electron beam of constant radius with variable mean electron velocity $v_0(x)$ by means of an inhomogeneous line, we arrive at the following equation [25]:

$$I'' + 3\frac{v_0'}{v_0} I' + \frac{\omega_p^2}{v_0^2} I = 0, \tag{3}$$

where $I' = dI/dx$; $I'' = d^2I/dx^2$; $I(x)$ is the amplitude of the AC component of the convection current in the cross section x (the factor $\left(i\omega t - i\int_0^x \beta_e dx\right)$ is omitted, see, for example, [2]); $v_0(x)$ is the mean velocity of the electrons in the x direction (mm/sec); $\omega_p(x)$ the effective plasma frequency: $\omega_p(x) = \omega_0(x)\, \alpha_0(\beta_e l)$; $\omega_0^2 = eI_0/m\varepsilon_0 v_0(x) S$; $e/m = 1{,}76 \cdot 10^{11}$ (coulombs/kg); $\epsilon_0 = 8.85 \cdot 10^{-12}$ (f/m); $v_0(x)$ (m/sec) is the DC component of the electron velocity; I_0 the DC component of the current in the beam amperes); S the area of beam cross section (square meters); $\alpha_0(\beta_e b)$ is a known function of $\beta_e b$; $\beta_e b = \omega b/v_0(x)$; ω the angular frequency of the signal; b the radius of the cylindrical electron beam. The function $\alpha_0(\beta_e b)$ is approximated by the expression (see, for example [26])

$$\alpha_0^2(\xi) = \begin{cases} (-0.146\xi^2 + 0.596\xi + 0.121)^2 & \text{for } 0.1 < \xi < 1.59, \\ 1 - \exp(-0.42\xi) & \text{for } 1.59 < \xi < \infty. \end{cases} \tag{4}$$

Let the cross section of the beam vary with the coordinate x; in this case S = S(x). If the variation of the cross section with the coordinate is slow, so that no higher modes are produced in the beam (that is, no spatial harmonics), the alternating current will as before be described by Eq. (3), but in the calculation of ω_p we must put $S = \pi b_0^2 \varphi^2(x)$ where $\varphi(x) = 1 + \Delta \sin \beta k(x)\, x$ (see formula (1)).

We can consider two types of solutions, I_1 and I_2, of Eq. (3) in accordance with the different boundary conditions on the virtual cathode:

1) $I_1(0) \neq 0$, $I_1'(0) = 0$ — excitation by current at the potential minimum,

2) $I_2(0) = 0$, $I_2'(0) \neq 0$ — excitation by velocity* at the potential minimum.

These two types of solutions are assumed to be uncorrelated, and the current I_1 leads to a noise factor F_1 in the basence of the current I_2, while I_2 leads to a noise factor F_2 in the absence of I_1.

The total uncorrelated noise factor F is given by the equation

$$F - 1 = F_1 - 1 + F_2 - 1.$$

To calculate the noise factor it is necessary to integrate Eq. (3) subject to suitable initial conditions [2]. In accordance with the foregoing, but putting S = S(x) in (3) in accordance with (1) and substituting $v_0(x) = \sqrt{2eU(x)/m}$, where U(x) is the potential given by curves I and II of Fig. 1, we arrive at the equation

$$I_{1,2}'' + 1.5\frac{U'}{U+0.07} I_{1,2}' + 9.5\cdot 10^{-3} \frac{I_{0n}\alpha_0^2(\xi)}{\pi b_0^2 (U+0.07)^{3/2}\varphi^2(x)} I_{1,2} = 0, \qquad (5)$$

where I_{0n} is the DC component of the current in the beam (in amperes); it assumes two values (n = 1, 2) in accordance with the two operating modes of TWT (Curves I and II of Fig. 1); $I_{1,2}$ is the AC component of the current, the indices 1 and 2 correspond to the two types of initial conditions; $\alpha_0^2(\xi)$ is given by formula (4); $\varphi(x) = 1 + \Delta_i \sin k(x)\, x\beta_i$;

U(x) is given by plots I and II in modes I and II, 350 v

$$x^2 q_{0m} = 2.33\, U^{3/2} + 1.84\, U \quad (0 \leq x \leq 0.1),$$
$$0.1 < x \leq 17,$$
$$17 \leq x \leq 29;$$

k(x) is given by formula (2); $b_0 = 5 \cdot 10^{-4}$ m; the figure 0.07 is the mean thermal energy of the electrons (in electron volts) at the potential minimum at a cathode temperature $T = 10^3$ °K (see, for example, [2]); $q_{0n} = I_{0n}/\pi b_0^2$, q_{0n} has the dimensionality of amperes per square meter.

In the foregoing calculations the ranges were $0 < \Delta < 0.4$; $0.5 < \beta < 4$. With the aid of expression (2) and Fig. 1 it can be verified that in the interval $0 < x < 29$ the inequality $\Lambda(x) \gg \lambda_e$ is satisfied, where $\Lambda(x) = 2\pi/k(x)\beta$ (mm); k is given by formula (2); λ_e (mm) $= 2\pi v_0(x)/\omega$; $v_0(x)$ is the mean velocity of the electrons in the cross section x (mm/sec); ω is the angular frequency. If this inequality is satisfied, we can neglect the spatial harmonics in the beam and use the equation for the inhomogeneous long line (5).

The values of the parameters Δ_i and β_i are listed in the table. Equation (5) must be integrated subject to the initial conditions

$$I_{1n}(0) \neq 0, \qquad (6)$$
$$I_{1n}'(0) = 0, \qquad (7)$$
$$I_{2n}(0) = 0, \qquad (8)$$
$$I_{2n}'(0) \neq 0. \qquad (9)$$

---

* We recall that in the linear theory $\tilde{I}' = -jI_0\omega \frac{\tilde{v}}{v_0^2} \cdot 10^{-3}$, where $\tilde{I}$ and $\tilde{v}$ are the AC components of the current and electron velocity; I_0 and v_0 are the DC components of the same quantity; ω the angular frequency; I and v have the dimensions of amperes and meters per second; ω is in sec^{-1}; I' = dI/dz (amp/mm).

No. variants	Δ_i	β_i	F_1-1	F_2-1	No. variants	Δ_i	β_i	F_1-1	F_2-1
1	0.15	0.5	0.36	18.9	*12*	0.3	0.2	0.29	18.1
2	0.15	1.0	0.38	16.4	*13*	0.3	0.5	0.15	54.7
3	0.15	1.15	0.65	11.6	*14*	0.3	1	0.48	24.5
4	0.2	1.0	0.39	18.7	*15*	0.3	1.5	0.63	10.8
5	0.3	1.0	0.47	23.2	*16*	0.3	2	0.62	12.6
6	0.4	1.0	0.83	28.4	*17*	0.15	0.5	72.3	682
7	0.15	0.2	0.58	10.8	*18*	0.15	1.0	29.1	418
8	0.15	0.4	0.44	13.3	*19*	0.15	1.5	96.4	1165
9	0.15	2	0.44	11.4	*20*	0.2	1	21.2	380
10	0.15	3	0.59	9.77	*21*	0.3	1	13.1	314
11	0.15	4	0.76	8.17	*22*	0.4	1	5.68	155

The values of (6) - (9) were obtained with the aid of the following formulas (see, for example, [2])

$$I_{1n}(0) = (2eI_{0n}\Delta f)^{1/2}, \qquad (10)$$

$$I'_{2n}(0) = \frac{I_{0n}\omega}{v_0^2(0)}\left[(4-\pi)\frac{e}{m}\frac{kT\Delta f}{I_{0n}}\right]^{1/2}\cdot 10^{-3}, \qquad (11)$$

where $e = 1.6\cdot 10^{-19}$ k; I_{0n} is the DC component of the current, in amperes; Δf the bandwidth of the receiver (cps); $\omega = 2\pi f$; f the working frequency of the amplifier (cps); $v_0(0)$ the beam velocity of the electrons (m/sec) at the potential minimum; $m = 9\cdot 10^{-31}$ kg; $k = 1.38\cdot 10^{-23}$ joule/deg; T the absolute temperature of the cathode. In these units, I_{1n} has the dimension of amperes and Γ_{2n} is in amperes per millimeter. By substituting in (10) and (11) the values $I_{01} = 2\cdot 10^{-4}$ amp, $I_{02} = 4\cdot 10^{-4}$ amp, $f = 3\cdot 10^9$ cps, $T = 10^{3\circ}$K, $\Delta f = 2\cdot 10^6$ cps and assuming also that $v_0^2(0)$ corresponds to an energy 0.07 ev, we obtain

$$q_{1.1}(0) = \frac{I_{1.1}(0)}{S_0} = 0.0144\ \frac{\text{amp}}{\text{m}^2}, \qquad q'_{2.1}(0) = \frac{I'_{2.1}(0)}{S_0} = 0.9\ \frac{\text{amp}}{\text{m}^2\cdot\text{mm}};$$

$$q_{1.2}(0) = \frac{I_{1.2}(0)}{S_0} = 0.0204\ \frac{\text{amp}}{\text{m}^2}; \qquad q'_{2.2}(0) = \frac{I'_{2.2}(0)}{S_0} = 1{,}275\ \frac{\text{amp}}{\text{m}^2\cdot\text{mm}},$$

where $S_0 = \pi b_0^2$; $b_0 = 5\cdot 10^{-4}$ meters.

We note first of all that when $x < 0.1$ mm the function $\varphi(x)$ entering into Equation (5) is approximately equal to 1, that is, the modulation of the beam cross section does not affect the behavior of the functions $I_{1,2}$ near the cathode. Equation (5) is a generalization of the Mathieu-Hill equation. It was integrated subject to initial conditions (6) - (9) for a series of values of the parameters Δ_1 and β_1 using a high-speed electronic computer. The solution was carried out for two operating modes of the TWT, for which the potential distribution U(x) is given by Curve I or II (Fig. 1). The function U(x) was specified here graphically, and was approximated by a sequence of third-degree polynomials. In the initial portion, the algebraic equation for the function U(x) was replaced by a differential equation.

The noise factors were automatically calculated after the integration of this equation by means of a program.

Figs. 2 — 7 show plots of the functions $q_1(x) = I_1(x)/S_0$ (amp/m^2), and $q_2(x) = I_2(x)/S_0$ (amp/m^2), where $S_0 = \pi b_0^2$ (m^2), for several values of the parameters Δ_1 and β_1, listed in the table. From these plots we can draw the following conclusions.

1. By comparing the course of the curves $q_1(x)$ and $q_2(x)$ in the presence of modulation, $0 < \Delta < 0.4$, with the case when there is no modulation, $\Delta = 0$ (this last case is considered in the article [2]), we conclude that these curves have qualitatively the same behavior.

2. On entering the helix ($x = L = 29$), the inequalities $|q_1(L)| < |q_2(L)|$ and $|q'_1(L)| < |q'_2(L)|$ are satisfied, and this leads to the inequality $F_1 < F_2$, where the uncorrelated values of the noise factors F_1 and F_2 are calculated with the aid of the formula

$$F_{1.2}(L) - 1 = Aq_{1.2}^2(L) - Bq_{1.2}(L)\,q'_{1.2}(L) + Eq'^2_{1.2}(L);$$

The coefficients A, B, and E depend of the geometry of the helix, on the beam parameters, and also on the operating mode of the TWT.

Analogous conclusions were made earlier in [2] for a TWT with a cylindrical electron beam of constant cross section.

For the sake of being definite, we choose two specific possible operating modes of the TWT

1. $I_{02} = 2\cdot 10^{-4}$ amp, $d_0 = 5\cdot 10^{-4}$ m. The distribution of the electric field is given by the curve I of Fig. 1. In this case $A = 3.08\cdot 10^8$, $B = 8.45\cdot 10^7$, $E = 1.25\cdot 10^9$.

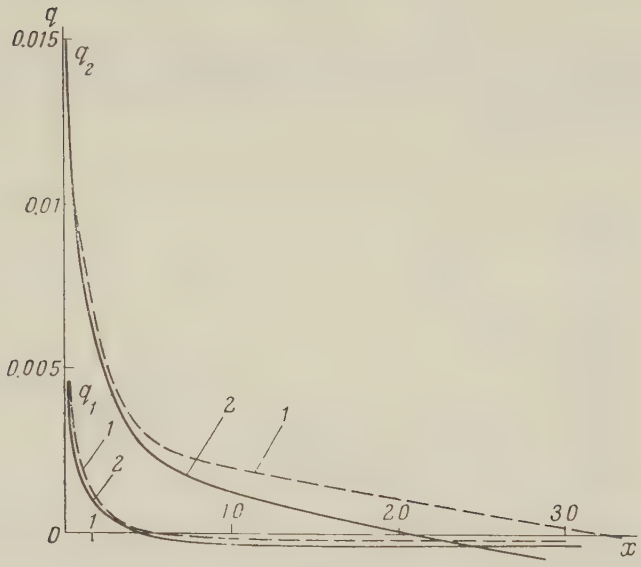


Figure 2. The functions $q_1(x)$ and $q_2(x)$:
1 — variant 1; 2 — variant 2

2. $I_{02} = 4 \cdot 10^{-4}$ amp, $d_0 = 5 \cdot 10^{-4}$ m; the distribution of the electrostatic field is given by Curve II of Fig. 1. In this case $A = 1.85 \cdot 10^7$, $B = 3.94 \cdot 10^8$, $E = 5.4 \cdot 10^9$.

We note that formula (12) must be regarded as approximate, since it takes into account the pulsation of the electron beam in the gun [through the quantities q(L) and q'(L)], but does not take into account the pulsation of the electron-beam cross section in the helix. An account of the pulsation of the beam in the helix complicates the solution of the problem and is not given here.

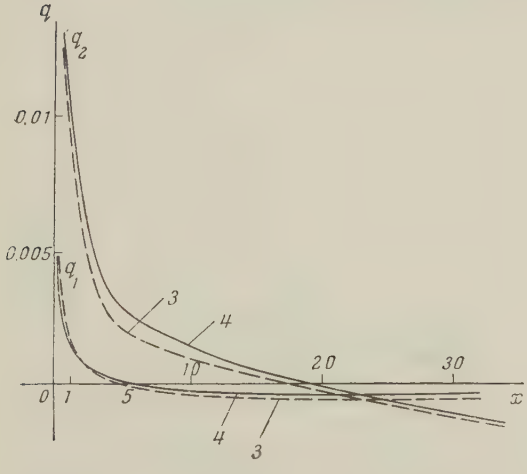


Figure 3. The functions $q_1(x)$ and $q_2(x)$:
3 — variant 3; 4 — variant 4

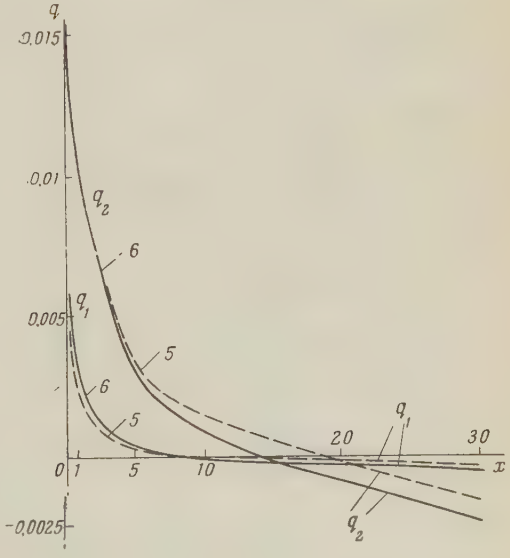


Figure 4. The functions $q_1(x)$ and $q_2(x)$:
5 — variant 5; 6 — variant 6

The calculated values of $F_1 - 1$ $F_2 - 1$, calculated by means of formulas (12) as functions of Δ_1 and β_1, are listed in the table. Figs. 8 and 9 show several curves $F_2(\Delta_1)$ and $F_2(\beta_1)$.

From the table and from Figs. 8 and 9 we can draw the following conclusions.

1. The inequality $F_2 \gg F_1$ is satisfied for these modes, and, consequently, in the case of uncorrelated excitation of noise waves at the potential minimum, the principal role is played by velocity fluctuations.

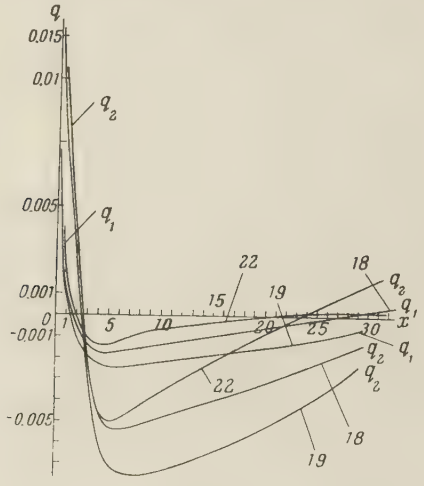


Figure 5. The functions $q_1(x)$ and $q_2(x)$

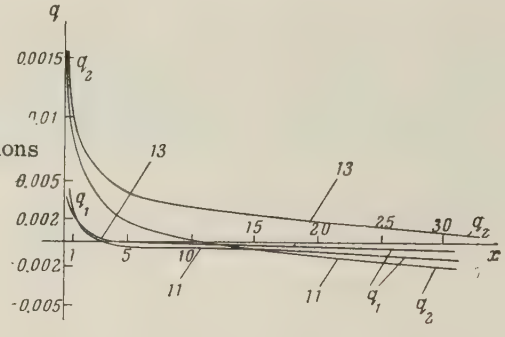


Figure 6. The functions $q_1(x)$ and $q_2(x)$

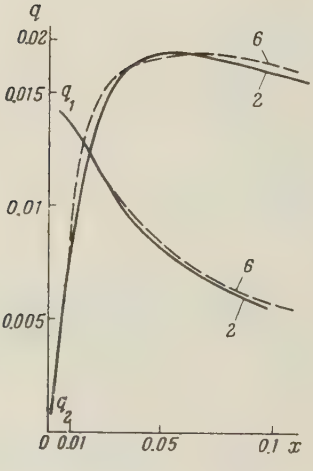


Figure 7. The functions $q_1(x)$ and $q_2(x)$

2. It follows from Fig. 8 that in mode I the plot of $F_2 - 1$ vs. β has a maximum at $\beta_1 \simeq 0.5$. The value of this maximum increases from 19 at $\Delta_1 = 0.15$ to 58 at $\Delta_1 = 0.3$. The curve $F_2(\beta)$ for mode II of the electron-gun operation is analogous in form.

3. It follows from Fig. 9 that $F_2 - 1$ in mode I monotonically increases with increasing Δ_1.

4. In all the plots given, the pulsation of the electron-beam cross section, $\Delta_1 \neq 0$, leads to an increase in the noise factor of the TWT compared with the case $\Delta_1 = 0$.

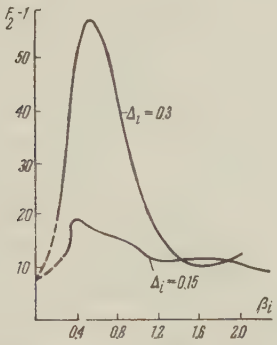


Figure 8. $F_2 - 1$ as a function of β_1 for $\Delta_1 = 0.15$ and $\Delta_1 = 0.3$ (mode I)

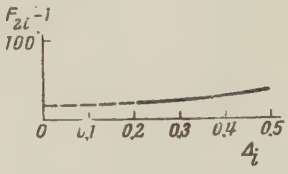


Figure 9. $F_{2i} - 1$ as a function of Δ_1 for $\beta_1 = 1$ (mode I), obtained for variants 4 — 10

Figs, 10a and b show the dependence of $F_2 - 1$ on $\theta = \beta_p l$, where $\beta_p = \frac{\omega_0}{v_0} a_0$ is the plasma wave number in the drift portion $U_0 = 350$ v in mode I (see Fig. 1); l the distance from the cathode to the enterance into the helix. In the mode considered ($U_0 = 350$ v, $I_{01} = 2 \cdot 10^{-4}$ amp, $b_0 = 5 \cdot 10^{-4}$ m, $\beta_p = 0.0315\ \text{mm}^{-1}$), the function $F_2(\theta)$ has been plotted for several values of the pulsation parameters Δ_1 and β_1 listed in the table (Curves 1 — 16).

The dashed curves on the same figures show the dependence of $F_2 - 1$ on θ for $\Delta = 0$, corresponding to a cylindrical beam with constant cross section $b = b_0$. The vertical dash-dot line corresponds to $l = L = 29$ mm; the intersections of this curve with the Curves 1 — 16 yield the values of $F_2 - 1$ listed in the table.

It follows from Figs. 10a and b that the pulsation of the beam cross section influences both the minimum value of the noise factor $F_{2\min}$ and the distance $l = l_{opt}$ from the cathode to the entrance into the helix, at which F_2 assumes a minimum value.

Fig. 11 shows the values of $F_{2\min} - 1$ as a function of β_1 for $\Delta_1 = 0.15$ and $\Delta_1 = 0.3$. These curves are similar in form. They have a maximum near $\beta \simeq 0.8$. With increasing Δ_1, the value of the function at the maximum increases.

Fig. 12 shows $F_{2\min} - 1$ as a function of Δ_1 for $\beta = 1$. This function has been plotted on the basis of Curves 4 — 6 of Fig. 10a. It follows from Fig. 12 that at fixed β an increase in depth of modulation Δ leads to a monotonic increase of $F_{2\min} - 1$ and consequently is undesirable from the point of view of reducing the noise factor F_2. It must be borne in mind, however, that at greater values of Δ_1 the long-line equation (3), which is the basis of the calculation, is no longer valid since it does not take into account the presence of spatial harmonics in the beam.

The curves given for $F_2(\theta) - 1$ indicate that the pulsation in the beam cross section almost always leads to an increase in the uncorrelated noise factor F_2, due to the fluctuations of the electron velocities on the virtual cathode.

Only in certain cases, with specially chosen Δ, β, and $\theta_{opt} = \beta_p l_{opt}$, can one obtain a slight decrease in $F_{2\min}$ compared with the minimum of the curve $F_2(l)$ $\Delta_i = 0$ for a cylindrical beam (see Curve 8 of Fig. 10b; $\Delta_i = 0.15$, $\beta_i = 0.4$).

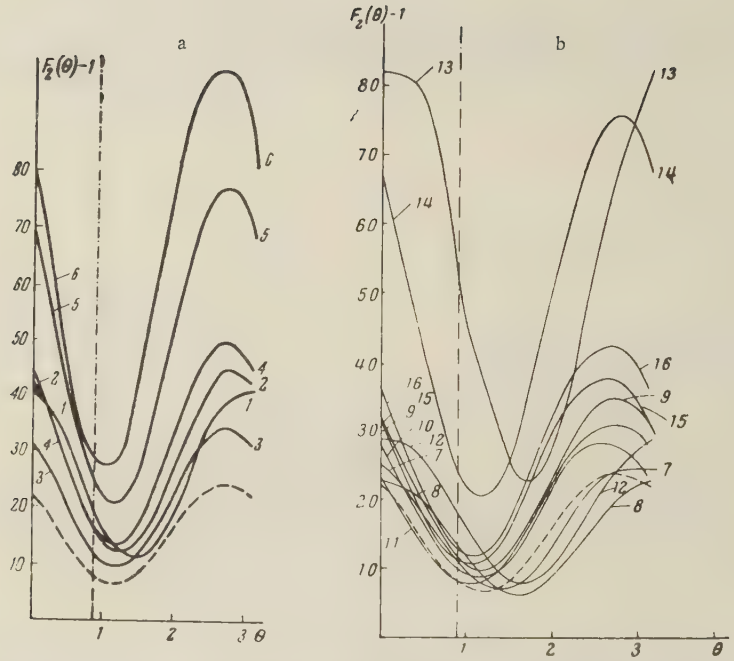


Figure 10. $F_2 - 1$ as a function of $\theta = \beta p l$ for the drift section in mode I.

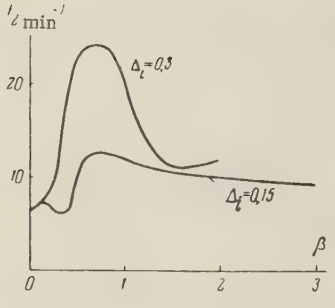


Figure 11. $F_{2min}(\theta_{opt}) - 1$ as a function of β for $\Delta_1 = 0.15$ and $\Delta_1 = 0.3$ in mode I (see Figure 1).

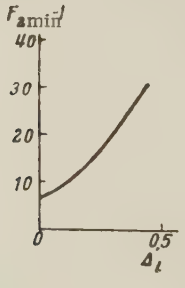


Figure 12. $F_{2min}(\theta_{opt}) - 1$ as a function of Δ_1 for $\beta_1 = 1$ of mode I (see Figure 1).

REFERENCES

1. A.S. Tager, Radiotekhnika i Elektronika, 1957, 2, 2, 222.
2. A.A. Zyuzin-Zinchenko, B.M. Lopukhin, V.M. Vasil'yev, Izvestiya Vuzov, Ministry of Higher Education USSR — Radio Engineering, 1959, 2, 5, 589.
3. V.N. Shevchik, G.N. Shvedov, ibid, 1959, 2, 5, 511.
4. Noises in Electron Devices, Ed. by L.D. Smullin, H.A. Haus, N.Y., 1959.
5. W. Kleen, K. Pöschl, Einfuhrung in die Mikrowellen Elektronik, II, Stuttgart, 1958.
6. J. Berghammer, S. Bloom, J. Appl. Phys., 1960, 31, 3, 454.
7. F.N.H. Robinson, J. Electronics and Control, 1958, 5, 2, 152.
8. P.K. Tien, J. Moshman, J. Appl. Phys., 1956, 27, 9, 1067.
9. M.E. Gertsenshteyn, Radiotekhnika i Elektronika, 1959, 4, 1, 146.
10. B.A. McIntosh, Canad. J. Phys., 1959, 37, 3, 265.
11. Saito Shigebumi, IRETrans., Electron Devices, 1958, 5, 4, 264.
12. M.R. Currie, D.C. Forster, Proc. I.R.E., 1958, 46, 3, 570.
13. M.R. Curie, D.C. Forster, J. Appl. Phys., 1959, 30, 1, 95.
14. M. Goulton, Proc. I.R.E., 1958, 46, 5, 911.
15. A.W. Shaw, A.E. Siegman, D.A. Watkins, Proc. I.R.E., 1959, 47, 2, 334.
16. W.M. Mueller, M.R. Currie, J. Appl. Phys., 1959, 30, 12, 1876.
17. M.E. Gertsenshteyn, Radiotekhnika i Elektronika, 1958, 3, 10, 1254.
18. R. Adler, Proc. I.R.E., 1958, 46, 6, 1300.
19. R. Adler, G. Hrbek, G. Wade, Proc. I.R.E., 1958, 46, 10, 1756.
20. R. Adler, Proc, I.R.E., 1959, 47, 10, 1713.
21. Curtis C. Johnson, J. Appl. Phys., 1960, 31, 2, 338.
22. B.M. Lopukhin, Radiotekhnika i Elektronika, 1961, 6, 5, 683.
23. R.C. Knechtly, W.R. Beam, Onde électr., 1958, 38, 371, 101.
24. B.M. Tsarev, Design and Construction of Vacuum Tubes, GEI, 1952.
25. S. Bloom, R. Peter, RCA Rev., 1954, 15, 1, 95.
26. D.A. Watkins, Proc. I.R.E., 1952, 40, 1, 165.
27. S. Bloom, RCA Rev., 1955, 16, 2, 179.

Physics Faculty,
V.M. Lomonosov Moscow State University
Chair of Radio Engineering.

Received by the editors 22 December 1960

MOLECULAR OSCILLATOR WITH TWO RESONATOR NATURAL FREQUENCIES WITHIN THE WIDTH OF THE EMISSION LINE

V. N. Lugovoy

We consider sinusoidal oscillations of a molecular oscillator which has two natural resonator frequencies within the width of the emission line of the material. It is established that in the case of identical resonator modes (that is, modes with equal Q and with equal coupling coefficient with the material) there is one stable mode of sinusoidal oscillations, if the detuning of the natural frequencies of the resonator does not exceed a critical value ($\Delta\omega_{cr}/\omega = 1/Q$, where Q pertains to the oscillation mode). If this detuning is greater than critical, there are two stable modes of sinusoidal oscillations with different frequencies. The oscillation frequency of a molecular oscillator is determined in this case by the initial conditions.

INTRODUCTION

At the present time great interest is evidenced in the problem of producing a molecular oscillator, whose oscillations lie in the infrared and optical regions [1,2,3]. The dimensions of the resonator in such an oscillator must exceed greatly the wavelength, and therefore many natural frequencies of the resonator will lie in the emission line width of the material. In order to understand the operation of the oscillator in this case, it is necessary to solve the problem of steady-state oscillations for at least two resonator modes. The present paper is devoted to this question.

1. ABBREVIATED EQUATIONS

The general equations for a molecular oscillator were derived in [4,5]. If two modes exist in the resonator, these equations yield

$$\begin{aligned}
&\ddot{x} + \mu_1\dot{x} + \varepsilon_1 x = -\gamma_1\ddot{v},\\
&\ddot{y} + \mu_2\dot{y} + \varepsilon_2 y = -\gamma_2\ddot{v},\\
&\ddot{v} + \mu_3\dot{v} + \varepsilon_3 v = \mu_3\left(-R_0 + \frac{w}{2}\right)(x+y),\\
&\dot{w} + \mu_3\frac{w}{2} = -\mu_3(x+y)\left(\dot{v} + \mu_3\frac{v}{2}\right).
\end{aligned} \tag{1}$$

These equations are in dimensionless quantities. The sum x + y describes the total electric field intensity inside the resonator; v describes the polarization of the material filling the resonator (for example, a beam); w gives the time dependence of the number of "radiating" molecules. The parameters contained in (1) are: μ_1 and μ_2 the damping coefficients of the resonator mode; μ_3 the emission line width of the material; ε_1, ε_2, and ε_3 the squares of the natural frequencies of the resonator and the material line; γ_1 and γ_2 the coupling coefficients of these modes with the substance; R_0 the initial number of active molecules. We shall assume that the dimensionless time is chosen such as to make $\varepsilon_3 = 1$. By making the change of variables $v = \mu_1 z$ and $w = \mu_1 a$ in Equation (1), we transform this equation to

$$\begin{aligned}
&\ddot{x} + \mu_1\dot{x} + \varepsilon_1 x = -\gamma_1\mu_1\ddot{z},\\
&\ddot{y} + \mu_2\dot{y} + \varepsilon_2 y = -\gamma_2\mu_1\ddot{z},\\
&\ddot{z} + \mu_3\dot{z} + z = \mu_3\left(-f_0 + \frac{a}{2}\right)(x+y),\\
&\dot{a} + \mu_3\frac{a}{2} = -\mu_3(x+y)\left(\dot{z} + \mu_3\frac{z}{2}\right),
\end{aligned} \tag{2}$$

where $f_0 = R_0/\mu_1$.

We shall assume henceforth the following conditions of practical interest to be fulfilled:

$$\begin{gathered}\mu_1, \mu_2 \ll \mu_3 \ll 1, \\ |\varepsilon_1 - 1|, \quad |\varepsilon_2 - 1| \ll \mu_3,\end{gathered} \tag{3}$$

which mean that the width of the resonance curve of each resonator mode is much less than the width of the emission line of the material; the width of the emission line is much less than its central frequency; both natural modes of the resonator are much closer to the center of the line than to its "edges." Assume, to be specific, that $\mu_1 \geq \mu_2$. We shall consider the coefficients γ_1, γ_2, and f_0 to be constants independent of these small parameters. Then the system (2) is close to a system of equations having general solutions

$$\begin{gathered}x = x_1 \cos t + x_2 \sin t, \\ y = y_1 \cos t + y_2 \sin t, \\ z = z_1 \cos t + z_2 \sin t, \\ a = \text{const}.\end{gathered} \tag{4}$$

Consequently to solve the problem we shall use the van der Pol method, assuming the constants in (4) to be "slowly varying amplitudes." We put $\xi = \dot{x}$, $\eta = \dot{y}$, $\zeta = \dot{z}$. Assuming small values of $z = -z$ and using (2), we obtain

$$\begin{aligned}&\dot{x} = \xi, \\ &\dot{\xi} = -\omega^2 x - \mu_1 \xi - (\varepsilon_1 - \omega^2)\, x + \gamma_1 \mu_1 z, \\ &\dot{y} = \eta, \\ &\dot{\eta} = -\omega^2 y - \mu_2 \eta - (\varepsilon_2 - \omega^2)\, y + \gamma_2 \mu_1 z, \\ &\dot{z} = \zeta, \\ &\dot{\zeta} = -z - \mu_3 \zeta + \mu_3 \left(-f_0 + \frac{a}{2}\right)(x + y), \\ &\dot{a} = -\mu_3 \frac{a}{2} - \mu_3 \zeta\,(x + y).\end{aligned} \tag{5}$$

In these expressions we introduced the frequency ω. We shall assume that

$$|\omega^2 - 1| \ll \mu_3, \tag{6}$$

then

$$|\varepsilon_1 - \omega^2|, \quad |\varepsilon_2 - \omega^2| \ll \mu_3. \tag{7}$$

Let further

$$\begin{aligned}x &= x_1(t) \cos \omega t + x_2(t) \sin \omega t, \\ \xi &= -\omega x_1(t) \sin \omega t + \omega x_2(t) \cos \omega t, \\ y &= y_1(t) \cos \omega t + y_2(t) \sin \omega t, \\ \eta &= -\omega y_1(t) \sin \omega t + \omega y_2(t) \cos \omega t, \\ z &= z_1(t) \cos \omega t + z_2(t) \sin \omega t, \\ \zeta &= -\omega z_1(t) \sin \omega t + \omega z_2(t) \cos \omega t.\end{aligned} \tag{8}$$

Here $x_1, x_2, y_1, y_2, z_1, z_2$ are new unknown functions. We substitute (8) in (5). Taking (3) and (6) into account and averaging over the explicitly expressed time,

we obtain

$$\begin{aligned}
x_1' &= -x_1 + \zeta_1 x_2 - \gamma_1 z_2,\\
x_2' &= -\zeta_1 x_1 - x_2 + \gamma_1 z_1,\\
y_1' &= -\zeta_3 y_1 - \zeta_2 y_2 - \gamma_2 z_2,\\
y_2' &= \zeta_2 y_1 - \zeta_3 y_2 + \gamma_2 z_1,\\
\frac{\mu_1}{\mu_3} z_1' &= -z_1 + \left(f_0 - \frac{a}{2}\right)(x_2 + y_2),\\
\frac{\mu_1}{\mu_3} z_2' &= -z_2 - \left(f_0 - \frac{a}{2}\right)(x_1 + y_1),\\
\frac{\mu_1}{\mu_3} a' &= -a + (x_2 + y_2)\, z_1 - (x_1 + y_1)\, z_2.
\end{aligned} \tag{9}$$

We put here $\zeta_1 = (\varepsilon_1 - \omega^2)/\mu_1$, $\zeta_2 = (\omega^2 - \varepsilon_2)/\mu_1$, $\zeta_3 = \mu_2/\mu_1$; the prime denotes the derivative with respect to the slow time τ:

$$\tau = \frac{1}{2}\mu_1 t. \tag{10}$$

It is easy to see that by virtue of (3) and (7) the equations (9) express a problem of the type

$$u' = G(u, v), \quad \beta v' = F(u, v), \tag{11}$$

where $u = \{x_1, x_2, y_1, y_2\}$; $v = \{z_1, z_2, a\}$; β is a small parameter. It is known that the phase space breaks up in this case into regions of "fast" and "slow" motions [6]. The fast motion is determined by the last three equations of the system (9) (it is necessary to assume u = const here). These equations comprise a linear system. Its characteristic equation can be readily seen to have the form

$$\left(1 + \frac{\mu_1}{\mu_3}\lambda\right)\left\{\left(1 + \frac{\mu_1}{\mu_3}\lambda\right)^2 + \frac{(x_1+y_1)^2 + (x_2+y_2)^2}{2}\right\} = 0, \tag{12}$$

hence

$$\operatorname{Re}\lambda_{1,2,3} = -\mu_3/\mu_1. \tag{13}$$

Taking in addition account of the fact that the relation $F(v, v) = 0$ in this case implies that to each u there corresponds only one value of v, we conclude that regardless of the initial data, the system considered will arrive by fast motion in the region of slow motions and will stay there for all subsequent instants of time. The equations of motion of this system in this region will be

$$u' = G(u, v), \quad 0 = F(u, v). \tag{14}$$

Expressing v in terms of u and substituting this expression in G(u, v) we obtain

$$\begin{aligned}
x_1' &= -x_1 + \zeta_1 x_2 + \gamma_1 f_0 \frac{x_1 + y_1}{1 + (1/2)\left[(x_1+y_1)^2 + (x_2+y_2)^2\right]},\\
x_2' &= -\zeta_1 x_1 - x_2 + \gamma_1 f_0 \frac{x_2 + y_2}{1 + (1/2)\left[(x_1+y_1)^2 + (x_2+y_2)^2\right]},\\
y_1' &= -\zeta_3 y_1 - \zeta_2 y_2 + \gamma_2 f_0 \frac{x_1 + y_1}{1 + (1/2)\left[(x_1+y_1)^2 + (x_2+y_2)^2\right]},\\
y_2' &= \zeta_2 y_1 - \zeta_3 y_2 + \gamma_2 f_0 \frac{x_2 + y_2}{1 + (1/2)\left[(x_1+y_1)^2 + (x_2+y_2)^2\right]}.
\end{aligned} \tag{15}$$

Equations (15) determine in particular all the limit cycles of the initial equations and their stability. According to (8), the equilibrium position of the system (15) corresponds to the sinusoidal oscillations of the molecular oscillator, and, conversely, any sinusoidal mode [whose frequency satisfies conditions (6)] can be obtained as an equilibrium position of this system for a certain value of ω.

2. SINUSOIDAL OSCILLATIONS

We shall henceforth consider the equilibrium positions $\tilde{x}_1$, x_2, $\tilde{y}_1$, $\tilde{y}_2$ of equations (15) and their stability. With respect to the equilibrium positions, the system (15) is equivalent [subject to conditions (3) and (6)] to the equations

$$\begin{aligned} \ddot{x} + \mu_1\dot{x} + \varepsilon_1 x &= -\gamma_1\mu_1\ddot{z}, \\ \ddot{y} + \mu_2\dot{y} + \varepsilon_2 y &= -\gamma_2\mu_1\ddot{z}, \\ \ddot{z} + \mu_3\dot{z} + z &= -\mu_3 (f_0 + \overline{(x+y)\dot{z}})\,(x+y), \end{aligned} \tag{16}$$

where

$$\begin{aligned} x &= \tilde{x}e^{-i\omega t} + \tilde{x}^* e^{i\omega t}; \\ y &= \tilde{y}e^{-i\omega t} + \tilde{y}^* e^{i\omega t}; \\ z &= \tilde{z}e^{-i\omega t} + \tilde{z}^* e^{i\omega t}; \\ \tilde{x} = \frac{\tilde{x}_1 + i\tilde{x}_2}{2}\ ;\quad \tilde{y} &= \frac{\tilde{y}_1 + i\tilde{y}_2}{2}\ ;\quad \tilde{z} = \frac{\tilde{z}_1 + i\tilde{z}_2}{2}\ ; \end{aligned} \tag{17}$$

The straight bar denotes averaging over the period. This can be verified directly (it becomes clear if the expressions (17) are substituted into the fourth equation of the system (2), from which it follows when $\mu_3 \ll 1$ that $a = -2\overline{(x+y)\dot{z}}$). Substituting (17) into (16) and recognizing that $f_0 + \overline{(x+y)\dot{z}}$ is a real constant from the compatibility condition, we obtain

$$\varkappa_3 = B\left(\frac{\gamma_1}{\varkappa_1} + \frac{\gamma_2}{\varkappa_2}\right), \tag{18}$$

where $\varkappa_k = -\omega^2 - i\omega\mu_k + \varepsilon_k$ $(k = 1, 2, 3;\ \varepsilon_3 = 1)$; B is a real number. Hence,

$$\frac{\operatorname{Re}\varkappa_3}{\operatorname{Im}\varkappa_3} = \frac{\operatorname{Re}\left(\frac{\gamma_1}{\varkappa_1} + \frac{\gamma_2}{\varkappa_2}\right)}{\operatorname{Im}\left(\frac{\gamma_1}{\varkappa_1} + \frac{\gamma_2}{\varkappa_2}\right)}. \tag{19}$$

Relation (19) is an algebraic equation for ω. Taking (3) into account, we can represent it in the form (see the appendix)

$$\begin{aligned} \Phi(v) = (\gamma_1 + \gamma_2)\,v^3 - (2\gamma_1 + \gamma_2)\,\Delta\varepsilon v^2 + \\ + (\gamma_2\mu_1^2 + \gamma_1\mu_2^2 + \gamma_1\Delta\varepsilon^2)\,v - \gamma_2\mu_1^2\Delta\varepsilon = 0, \end{aligned} \tag{20}$$

where $v = \omega^2 - \varepsilon_1$; $\Delta\varepsilon = \varepsilon_2 - \varepsilon_1$; $\Delta\varepsilon^2 = (\varepsilon_2 - \varepsilon_1)^2$. Each real root of the polynomial $\Phi(v)$ determines the frequency of possible sinusoidal oscillations.

The general expression for the amplitudes of these oscillations are cumbersome and are of no interest for the present analysis. We shall therefore not give them here. We note only that the dependence of the amplitude of the oscillations of the electric field on the initial number of active molecules is determined by the relation

$$f_0 = f_{0\,\mathrm{thr}}\,(1 + A^2), \tag{21}$$

where $A^2 = [(\tilde{x}_1 + \tilde{y}_1)^2 + (\tilde{x}_2 + \tilde{y}_2)^2]/2$. In this case the inequality $f_0 > f_{0\,\mathrm{thr}}$ and condition (20) are necessary and sufficient for the existence of nontrivial equilibrium positions for Equations (15).

The number of real roots of the polynomial $\Phi(v)$ is determined by the sign of its discriminant D (when $D > 0$ there is one real root). This discriminant has a simple form in the case of like coupling coefficients between the resonant oscillation modes and the substance $(\gamma_1 = \gamma_2 = \gamma)$;

$$D = \frac{1}{216}\left(\mu_1^2 + \mu_2^2 - \frac{1}{2}\Delta\varepsilon^2\right)^3 + \frac{\Delta\varepsilon^2}{64}\,(\mu_1^2 - \mu_2^2)^2. \tag{22}$$

The equation $D(\Delta\varepsilon^2) = 0$ has a unique positive solution for all μ_1 or μ_2, and determines the critical value of the detuning of the natural frequencies of the resonator. With this, $D(\Delta\varepsilon^2) > 0$ when $\Delta\varepsilon^2 < \Delta\varepsilon^2_{cr}$ and $D(\Delta\varepsilon^2) < 0$ when $\Delta\varepsilon^2 > \Delta\varepsilon^2_{cr}$.

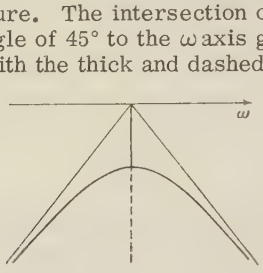
Let us consider several cases.

1). $\gamma_1 = \gamma_2 = \gamma$, $\mu_1 = \mu_2 = \mu$ (like oscillation modes).

We have $\Phi(v) = \gamma(2v - \Delta\varepsilon)(v - v_2)(v - v_3)$, where $v_{2,3} = \frac{\Delta\varepsilon}{2} \pm \sqrt{\frac{\Delta\varepsilon^2}{4} - \mu^2}$. Therefore, $\Delta\varepsilon_{cr}^2 = 4\mu^2$ or

$$\Delta\omega_{cr}/\omega = 1/Q, \tag{23}$$

Q is the figure of merit at the modes under consideration. The location of the possible oscillation frequencies can be determined for any value of the natural-frequency detuning with the aid of the figure. The intersection of a line parallel to the ω axis with the thin straight lines making an angle of 45° to the ω axis gives the location of the natural frequencies, while the intersection with the thick and dashed lines gives the location of the possible oscillation frequencies.



2. $\mu_2 \ll \mu_1$.

In this case $\Phi(v) = 2\gamma(v - \Delta\varepsilon)(v - v_2)(v - v_3)$, where $v_{2,3} = \frac{\Delta\varepsilon}{4} \pm (1/4)\sqrt{\Delta\varepsilon^2 - 8\mu_1^2}$ (we put for simplicity $\gamma_1 = \gamma_2 = \gamma$).

$$\Delta\omega_{cr}/\omega = \sqrt{2}/Q_1, \tag{24}$$

Q_1 refers to the first type of oscillations.

3. $\gamma_2 = 0$.

The polynomial $\Phi(v)$ has one real root $v = 0$, corresponding to the only possible oscillation frequency coinciding with the natural frequency of the first oscillation mode. Analogously, when $\gamma_1 = 0$ the polynomial $\Phi(v)$ has one real root, corresponding to oscillations at the second natural frequency.

3. STABILITY OF SINUSOIDAL OSCILLATIONS

Only the stable equilibrium positions of the system (15) correspond to practically realizable sinusoidal oscillations of a molecular oscillator. The stability of the harmonic oscillations of an oscillator with one natural resonator frequency, within the limits of the line width, was considered in [7,8,9]. In view of the fact that the initial equations are autonomous, any equilibrium position of this system in its phase space belongs to a closed phase trajectory (the orbit of the limit cycle). On the other hand, each point of the orbit is an equilibrium position of Equations (15). Consequently the characteristic polynomial of these equations, linearized in the vicinity of any equilibrium position, will have a single zero root. This root corresponds to "indifferent" motions along the limit-cycle orbit. The remaining three roots determine the orbit stability of the cycle, that is, the stability with respect to small deviations from the orbit.

We confine ourselves below to the case of identical resonant modes ($\gamma_1 = \gamma_2 = \gamma$, $\mu_1 = \mu_2 =$ We shall designate as "central" oscillations those for which $\zeta_1 = \zeta_2 = \zeta$ (see the figure), and as "sideband" oscillations those for which $\zeta_1 \zeta_2 = 1$. Using (16) and (6) we find

$$\gamma f_0 = \chi(1 + A^2), \tag{25}$$

where $\chi = (1 + \zeta^2)/2$ for central oscillations and $\chi = 1$ for "sideband" oscillations. From (25) it is seen that $f_{0thr} = \chi/\gamma$ [see (21)]. Linearizing the system (15) as usual and taking (25) into account, we obtain the characteristic equation for the central oscillations:

$$(d^2 + \zeta^2 - \varkappa^2)^2 - 4(\chi - \varkappa)^2(d - \varkappa)^2 = 0, \tag{26}$$

where $d = \varkappa - 1 - \lambda$; $\varkappa = \chi/(1 + A^2)$. This equation can be readily solved (with respect to λ) Its nonzero roots have the form

$$\lambda_1 = -(1 - \zeta^2),$$

$$\lambda_{2,3} = -\frac{(1-\zeta^2)}{2(1+A^2)}\left\{1 \pm (1 - A^2)\sqrt{1 - \frac{16\zeta^2 A^2}{(1-\zeta^2)^2(1-A^2)^2}}\right\} - \frac{(3+\zeta^2)A^2}{2(1+A^2)}. \tag{27}$$

It is easy to verify that $\mathrm{Re}\,\lambda_{2,3} < 0$ when $\zeta^2 < 1$ for all values of A^2. Therefore the equilibrium positions of the system (15), corresponding to the central oscillations, are stable if the detuning of the natural frequencies of the oscillator is less than critical, and are unstable if this detuning is greater than critical. In the figure, the unstable central oscillations correspond to the dashed line.

Analogously, we obtain a characteristic equation for the sideband oscillations

$$\lambda\left\{\lambda^3 + \frac{4A^2}{1+A^2}\lambda^2 + 4\left(\bar{\zeta}^2 - \frac{1}{1+A^2}\right)\lambda + \frac{8A^2}{1+A^2}(\bar{\zeta}^2 - 1)\right\} = 0. \tag{28}$$

Here $\bar{\zeta} = (\zeta_1 + \zeta_2)/2$. Recognizing that $\bar{\zeta}^2 > 1$ (when $\zeta_1 \neq \zeta_2$) and using the Routh-Hurwitz conditions, we can readily verify that $\mathrm{Re}\,\lambda_{1,2,3} < 0$ [$\lambda_{1,2,3}$ are the nonzero roots of Equation (28)]. Thus, the equilibrium positions corresponding to the sideband oscillations are always stable.

Equations (15) were obtained under the assumption that the solutions of the system (2) are close to solutions of the form (4). If the initial number of active molecules is sufficiently high (so that $\mu_3 f_0 \gtrsim 1$), then it can be readily seen that the solutions of the system (2) [subject to conditions (3)] will not be close to the form (4). This however is not essential in the problem of the limit cycles themselves (which determine the sinusoidal oscillations), since the expressions (16) can be obtained under these conditions for all f_0 directly from the system (2). However, at sufficiently high electric-field oscillation amplitudes it may turn out that the stability of these cycles is not equivalent to the stability of the equilibrium positions of the system (15).

Thus, if the initial number of active molecules is not too large, then the frequencies of all the stable oscillations of the molecular oscillator (in the case of identical resonator modes) are represented on the figure by the heavy continuous lines. Thus, if the detuning of the natural frequency of the generator is less than critical (23), then the oscillation frequency is unique; in the opposite case this frequency is determined by the initial conditions and is one of the two shown in the figure. If the detuning is considerably greater than critical, then a stable sinusoidal oscillation mode should be observed at each of the natural frequencies.

APPENDIX

It can be verified directly that Equation (19) is equivalent to

$$\begin{aligned} L(\omega) = &-\mu_3(\gamma_1+\gamma_2)(u-\varepsilon_1)^3 - (\gamma_1\mu_1+\gamma_2\mu_2)(u-\varepsilon_1)^2(u-1) + \\ &+ \Delta\varepsilon\,[\mu_3(2\gamma_1+\gamma_2)(u-\varepsilon_1)^2 + 2\gamma_1\mu_1(u-\varepsilon_1)(u-1)] - \mu_3\,[u(\gamma_2\mu_1^2+\gamma_1\mu_2^2)+\gamma_1\Delta\varepsilon^2]\times \\ &\times(u-\varepsilon_1) - [u(\gamma_2\mu_1^2\mu_2+\gamma_1\mu_2^2\mu_1)+\gamma_1\mu_1\Delta\varepsilon^2](u-1) + u\gamma_2\mu_3\mu_1^2\Delta\varepsilon = 0, \end{aligned} \tag{29}$$

where $u = \omega^2$. Replacing u by unity in (29) [wherever u does not enter into the difference $(u - \epsilon_1)$ or $(u-1)$] and using (3) we obtain the approximate expression $\Phi(v) = -L(\omega)/\mu_3$, which coincides with (20). It can be shown that Equation (20), as a relation defining the frequencies of the possible sinusoidal oscillations, is valid for any placement of the natural frequencies of the resonator within the limits of the emission line width (that is, when $|\varepsilon_1 - 1|, |\varepsilon_2 - 1| \lesssim \mu_3$).

The author is deeply grateful to N.G. Basov, F.V. Bunkin, and A.N. Orayevskiy for formulating the problem and for useful discussions during the work.

REFERENCES

1. A.L. Schawlow, C.H. Townes, Phys. Rev., 1958, 112, 6, 1940.
2. N.G. Basov, O.N. Krokhin, Yu.M. Popov, Uspekhi Fiz. Nauk, 1960, 72, 2, 1961.
3. A.M. Prokhorov, Zh. Eksp. Teoret. Fiz, 1958, 34, 6, 1958.
4. A.N. Orayevskiy, Radiotekhnika i Elektronika, 1959, 4, 4, 718.
5. V.M. Fayn, Zh. Eksp. Teoret. Fiz. 1957, 33, 4, 945.
6. A.A. Andronov, A.A. Vitt, S.E. Khaykin, Theory of Oscillations, GIFML, 1959.
7. V.S. Troitskiy, Radiotekhnika i Elektronika, 1958, 3, 10, 1298.
8. H.J. Haldre, R.V. Khokhlov, Izvestiya Vuzov, Ministry of Higher Education USSR—Radiophysics, 1958, 1, 5-6, 60.
9. A.S. Gurtovnik, ibid. 1958, 1, 5-6, 83.

P.N. Lebedev Physics Institute,
Academy of Sciences USSR

Submitted to Editors 27 December 1960

NEW PRINCIPLE OF MICROWAVE AMPLIFICATION WITH FERRITES

V. A. Fabrikov

A new principle is considered for the amplification of microwaves with ferrites; this method is connected with nonlinear gyromagnetic effects of the nutational type.

1. Nonlinear gyromagnetic effects in ferrites at microwave frequencies can be subdivided, in accordance with the principles of their utilization, into two groups, one connected with the frequency characteristic of the precessional motion of the spin moment, and the other connected with the angular characteristic of this motion. The first group of effects, used to develop different types of parametric ferrite microwave amplifiers, has been discussed in the literature in detail (see, for example, [1]). Effects of the second type, called nutational effects, have been studied to a much lesser extent.

It was shown earlier that nonlinear phenomena of the nutational type can be used to amplify electromagnetic signals at intermediate frequencies [2] and to increase the efficiency of microwave mixers [3]. Reference [3] points out also the possibility of obtaining a new effect—power amplification of one of the mixed microwave signals, at the expense of another stronger signal. This effect, which can be regarded as a new principle of microwave amplification with ferrites (its distinguishing feature is that the amplification takes place only in the presence of magnetic losses in the medium), is analyzed in the present article. The analysis is based on solving the equations of motion of the gyromagnetic moments of the magnetized ferrite material under the influence of two transversely polarized microwave signals of near-equal frequency and the longitudinal alternating reaction field. The solution of the problem is sought in a coordinate system rotating in synchronism with the magnetic moment of the specimen, this moment being of constant magnitude.

2. We consider a ferrite specimen under the influence of a constant magnetic field $\vec{H}_0 = \vec{i}_z H_0$ sufficiently strong to produce magnetic saturation in the material. By $\vec{H}_0$ we shall mean the effective value of the field inside the specimen, taking into account the demagnetizing factors due to the specimen shape. We assume that along with the constant field $\vec{H}_0$ the specimen is acted upon also by a certain arbitrarily oriented small alternating field $\vec{h}$, that is, $\vec{H} = \vec{H}_0 + \vec{h}$ and $h \ll H_0$. Then the magnetic moment $\vec{M}$ of the specimen can be regarded as a whole and its equation of motion can be written in vector form as follows [4]:

$$\dot{\vec{M}} = \gamma [\vec{H}\vec{M}] - \vec{R}, \qquad (1)$$

where the term $\vec{R}$, which takes the losses into account, has components $R_x = M_x/T_2$ $R_y = M_y/T_2$ $R_z = (M_z - M_0)/T_1$ [5, 6]. Here $\gamma = 1.76 \cdot 10^7$ oe^{-1} sec^{-1} is the absolute value of the gyromagnetic ratio of the electron spin; T_1 and T_2 are the longitudinal and transverse relaxation times, characterizing the intensity of the damping processes in the material; M_0 is the equilibrium value of the magnetic moment in a constant field H_0. The dot over a symbol denotes differentiation with respect to the time t, and the brackets denote the vector products of the corresponding quantities.

The foregoing form of the equations of motion of the magnetic moment does not presuppose that any limitations are imposed whatever on T_1 or T_2, and the three scalar equations corresponding to the vector equation (1) are in general independent. In most practical applications however, the absolute magnitude of the magnetic moment M can apparently be regarded as unchanged under the influence of the magnetic field, so that we can put $M = M_0$. This condition is equivalent to the existence of a definite connection between T_1 and T_2 ($T_2 \simeq 2T_1$) and between the quantities m_x, m_y, and m_z, where $\vec{m} = \vec{M} - \vec{M}_0$. Only two of the three equations of motion are independent here, so that it becomes expedient to consider the problem geometrically in a rotating coordinate system which is stationary with respect to the magnetic moment $\vec{M}$.

We introduce along with the Cartesian system x, y, z an orthogonal system ξ, η, ζ with unit vectors $\vec{i}_\xi$, $\vec{i}_\eta$, $\vec{i}_\zeta$ fixed relative to the moving vector of the magnetic moment $\vec{M}$. The ζ

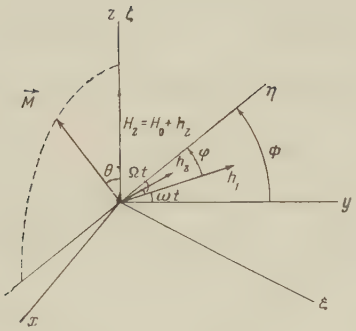


Figure 1. Schematic representation of the precessing magnetic moment M and the magnetic fields $H_Z = H_0 + h_Z$ and $h_x + ih_y = h_1 e^{i\omega t} + h_2 e^{i(\omega + \Omega)t}$ acting on it in a rotating system of coordinates ξ, η, ζ.

axis will be aligned with the constant magnetic field $\vec{i}_Z H_0$, and the η axis is located in the plane of the vectors $\vec{M}$ and $\vec{H}_0$ so that the angle between $\vec{M}$ and $\vec{i}_\eta$ exceeds 90° (Fig. 1). We denote by Φ the angle between the axes ξ and x (or η and y) and by Θ the angle between $\vec{M}$ and $\vec{i}_\zeta = \vec{i}_Z$. The angles are measured in the directions of the arrows shown in Fig. 1. It is obvious that $\Phi = \Phi(t)$. In the general case we also have $\Theta = \Theta(t)$.

Without imposing any additional limitations on the longitudinal component of the alternating field $\vec{h}$, we shall assume its transverse components to be harmonic and circularly polarized, setting

$$h_x = h_1 \cos \omega t, \qquad h_y = h_1 \sin \omega t$$

or in complex notation

$$h_x + ih_y = h_1 e^{i\omega t}, \tag{2}$$

where i is the imaginary unit. A consideration of only alternating fields with circular transverse polarization in the mathematical analysis of the phenomena that take place under conditions of ferromagnetic resonance, affects very little the generality of the results obtained. The form of notation in (2), which is convenient in describing fields with circular polarization, must not be confused with the symbolic representation of quantities that have a harmonic dependence on the time.

By confining ourselves to fields of circular polarization we can represent the transverse components of the field $\vec{h}$ by means of a vector $\vec{h}_1$ (in the plane xy) rotating in the stationary coordinate system with angular frequency ω. The angle $\varphi = \Phi - \omega t$, which defines the position of the vector $\vec{h}_1$ in the rotating system ξ, η, ζ can in the general case likewise not be regarded as constant.

We denote by ω' the instantaneous angular velocity of rotation of the vector $\vec{M}$ in the stationary frame. From geometrical considerations we can readily see that vectors $\vec{H}$, $\vec{M}$, and $\vec{\omega}'$ can be represented in the coordinate system ξ, η, and ζ in the form

$$\begin{gathered} \vec{H} = \vec{i}_\xi h_1 \sin \varphi + \vec{i}_\eta h_1 \cos \varphi + \vec{i}_\zeta (H_0 + h_z), \\ \vec{M} = -M(\vec{i}_\eta \sin \Theta - \vec{i}_\zeta \cos \Theta), \\ \vec{\omega}' = -\vec{i}_\xi \dot{\Theta} + \vec{i}_\zeta \dot{\Phi}. \end{gathered} \tag{3}$$

Here

$$\begin{gathered} \frac{\vec{M}}{M} = \frac{1}{M}[\vec{\omega}'\vec{M}] = \vec{i}_\xi \sin \Theta \dot{\Phi} - \vec{i}_\eta \cos \Theta \dot{\Theta} - \vec{i}_\zeta \sin \Theta \dot{\Theta}, \\ \frac{[\vec{H}\vec{M}]}{M} = \vec{i}_\xi \{h_1 \cos \Theta \cos \varphi + \sin \Theta (H_0 + h_z)\} - \vec{i}_\eta h_1 \cos \Theta \sin \varphi - \\ - \vec{i}_\zeta h_1 \sin \Theta \sin \varphi. \end{gathered} \tag{4}$$

Substituting (4) in (1) we arrive at a system of two nonlinear differential equations in the

independent variables φ and Θ [7]

$$\dot{\varphi} = (\gamma H_0 - \omega) + \gamma h_z + \gamma h_1 \operatorname{ctg} \Theta \cos \varphi, \qquad (5)$$
$$\dot{\Theta} = \gamma h_1 \sin \varphi - \frac{\operatorname{tg} \Theta}{T_2}.$$

The connection between the new variables φ, Θ and the old variables M_x, M_y, M_z is given by the relation

$$M_x + iM_y = -M \sin \Theta e^{i\Phi}, \qquad (6)$$
$$M_z = M \cos \Theta.$$

3. For what follows it is convenient to introduce a special symbol for the dimensionless parameter

$$x = (\omega - \gamma H_0) T_2 = \frac{H_{res} - H_0}{\Delta H}, \qquad (7)$$

which characterizes the proximity to the state of ferromagnetic resonance. Here $H_{res} = \omega/\gamma$ is the resonant value of the magnetizing field with respect to a signal of frequency ω, and $\Delta H = 1/\gamma T_2$ is the half-width of the ferromagnetic-resonance line.

Equations (5) can be readily solved if $h_z = 0$ and consequently $\dot{\varphi} = \dot{\Theta} = 0$. This solution in the new notation has the form

$$\operatorname{tg} \varphi_0 = \frac{1}{x},$$
$$\operatorname{tg} \Theta = \frac{1}{\sqrt{1 + x^2}} \frac{h_1}{\Delta H}. \qquad (8)$$

When $h_z \neq 0$, the exact solution of (5) is quite complicated. We can, however, find an approximate solution of these equations by regarding h_z as a small perturbation of the system, the unperturbed state of which is characterized by Equations (8).* Expanding the right halves of (5) in a Taylor series about the point φ_0, Θ_0 and discarding terms of second order and higher, we arrive at two linear differential equations in the quantities $\varphi_1 = \varphi - \varphi_0$ and $\Theta_1 = \Theta - \Theta_0$:

$$T_2 \dot{\varphi}_1 + \varphi_1 = \frac{h_z}{\Delta H} - x \sqrt{1 + x^2} \frac{\Delta H}{h_1} \Theta_1,$$
$$T_2 \dot{\Theta}_1 + \Theta_1 = \frac{x}{\sqrt{1 + x^2}} \frac{h_1}{\Delta H} \varphi_1. \qquad (9)$$

The quantity under the square root in (9) should, strictly speaking, be $1 + x^2 + \left(\frac{h_1}{\Delta H}\right)$ but we neglect the quantity $(h_1/\Delta H)^2$ by setting

$$\left(\frac{h_1}{\Delta H}\right)^2 \ll 1. \qquad (10)$$

For a specified h_z the solution of (9) is relatively simple and enables us to determine the reaction of the gyromagnetic medium to the effect of a weak modulating signal, described in the direction of the z-axis (in the case of a sinusoidal field h_z) by the complex magnetic susceptibility of the medium at the frequency of this field [7].

If the ferrite material is acted upon simultaneously by two transverse circularly polarized magnetic fields, that is,

$$h_x + ih_y = h_1 e^{i\omega t} + h_2 e^{i(\omega + \Omega)t}, \qquad (11)$$

then the expression $\vec{H}$ in the rotating frame assumes the more general form

$$\vec{H} = \vec{i}_\xi \{h_1 \sin \varphi - h_2 \sin (\Omega t - \varphi)\} + \vec{i}_\eta \{h_1 \cos \varphi + h_2 \cos (\Omega t - \varphi)\} + \vec{i}_\zeta (H_0 + h_z). \qquad (12)$$

* The perturbation can be regarded as small if $h_z \ll \Delta H$. In the appendix we consider the case when h_z is not necessarily small compared with Δh.

The initial phase shift is set equal to zero in Equations (11) and (12), for the sake of simplicity, this being immaterial. Considering only those cases in which $h_2 \ll h_1$, we can leave formulas (8) unchanged, and replace (9) by the more general equations

$$T_2\dot{\varphi}_1 + \varphi_1 = \frac{h_z}{\Delta H} - x\sqrt{1+x^2}\,\frac{\Delta H}{h_1}\,\theta_1 + \frac{h_2}{h_1}\sqrt{1+x^2}\cos(\Omega t - \varphi_0),$$
$$T_2\dot{\theta}_1 + \theta_1 = \frac{x}{\sqrt{1+x^2}}\,\frac{h_1}{\Delta H}\,\varphi_1 - \frac{h_2}{\Delta H}\sin(\Omega t - \varphi_0). \tag{13}$$

After determining the parameters φ_1 and θ_1 from (13), we can find expressions for the alternating components of the magnetic moment m_x, m_y, and m_z, using the formulas

$$m_x + im_y = -M\sin\theta e^{i\Phi} \simeq -M(\sin\theta_0 + \nu)\,e^{i(\omega t+\varphi_0)},$$
$$m_z = M(\cos\theta - \cos\theta_0) \simeq -M\sin\theta_0\theta_1, \tag{14}$$

where

$$\sin\theta_0 = \frac{1}{\sqrt{1+x^2}}\,\frac{h_1}{\Delta H};$$
$$\nu = \theta_1 + i\sin\theta_0\varphi_1. \tag{15}$$

Thus, we shall be interested henceforth in the variable quantities ν and θ_1. We rewrite (13) in terms of these quantities in the form

$$T_2\dot{\nu} + (1+ix)\,\nu = i\,\frac{1}{\sqrt{1+x^2}}\,\frac{h_1h_z}{(\Delta H)^2} + i\,\frac{h_1}{\Delta H}\,e^{i(\Omega t-\varphi_0)},$$
$$T_2^2\ddot{\theta}_1 + 2T_2\dot{\theta}_1 + (1+x^2)\,\theta_1 = \frac{x}{\sqrt{1+x^2}}\,\frac{h_1h_z}{(\Delta H)^2} + \tag{16}$$
$$+\frac{h_2}{2\Delta H}\{(x+i-\Omega T_2)\,e^{i(\Omega t-\varphi_0)} + (x-i-\Omega T_2)\,e^{-i(\Omega t-\varphi_0)}\},$$

which is most convenient for analysis. The right half of the last expression contains the sum of two complex-conjugate quantities.

4. So far we have imposed no limitations whatever on the variable quantity h_Z, contained in (16), except that it be small ($h_Z \ll H_0$), taking h_Z to mean an alternating component of the magnetic field $H_Z = H_0 = h_Z$, arbitrary in its spectral composition. This field determines the resonant characteristics of the material with respect to transversely polarized microwave fields at frequencies close to the natural frequency γH_Z of the material. The modulation of the magnetizing field H_Z can be produced by acting on the ferrite specimen with an external signal, the magnetic field of which is polarized in the z direction, but can also be produced by the reaction of the specimen or of a resonant circuit interacting with the specimen to the change in magnetization of the material in the direction of this axis.

In the present investigation we are interested in precisely this latter case, which was already considered earlier [8] from the point of view of the possibility of and the conditions for the occurrence of natural magnetization oscillations in a ferrite specimen, under the influence of a sufficiently powerful transversely polarized microwave signal. If the ferrite specimen is acted upon simultaneously by two transversely polarized microwave signals, one of which exceeds the other considerably in magnitude, the presence of the reaction field can lead, as shown below, to the occurrence of a nonlinear effect on top of the bias effect, namely the amplification of the weak microwave signal at the expense of the stronger signal. A further examination reduces to the analysis of the equations (14) and (16) for the particular case corresponding to the two of the most characteristic of the possible couplings between the quantities h_Z and θ_1 contained in (16).

Case 1. The field h_Z is the field due to the reaction of a resonant circuit interacting with the ferrite specimen (Figure 2). If the resonant circuit is tuned to a frequency Ω which coincides with the difference frequency between two microwave signals acting on the specimen, then by neglecting the higher harmonics of the oscillations m_Z we can represent the field of the reaction of the resonant circuit in the form [8]

$$h_z = -\eta\,\frac{L}{R}\,\dot{m}_z = \eta Q M\sin\theta_0\dot{\theta}_1T_2, \tag{17}$$

where Q is the figure of merit of the resonant circuit at a frequency $1/T_2$; η is the coefficient of coupling between the resonant circuit and the specimen, which is close to the filling coefficient.

Substituting (17) into (16) and introducing the additional notation

$$y = \Omega T_2, \quad a = \frac{M h_1^2}{(\Delta H)^3} \tag{18}$$

for the frequently encountered combinations of quantities, we arrive at the equations

$$\begin{aligned} T_2 \dot{\nu} + (1 + ix)\,\nu &= i\eta Q \frac{a}{1+x^2}\,\dot{\Theta}_1 T_2 + i\frac{h_2}{\Delta H} e^{i(\Omega t - \varphi_0)}, \\ T_2^2 \ddot{\Theta}_1 + \left(2 - \eta Q a \frac{x}{1+x^2}\right)\dot{\Theta}_1 T_2 + (1+x^2)\,\Theta_1 &= \\ = \frac{h_2}{2\Delta H}\{(x - y + i)\, e^{i(\Omega t - \varphi_0)} + (x - y - i)\, e^{-i(\Omega t - \varphi_0)}\}. & \end{aligned} \tag{19}$$

The solution of these equations (in the steady state) has the form

$$\begin{aligned} \Theta_1 = \frac{h_2}{2\Delta H}\Bigg\{ & \frac{x - y + i}{1 + x^2 - y^2 + 2iy\left(1 - \eta Q \frac{a}{2}\frac{x}{1+x^2}\right)} e^{i(\Omega t - \varphi_0)} + \\ & + \frac{x - y - i}{1 + x^2 - y^2 - 2iy\left(1 - \eta Q \frac{a}{2}\frac{x}{1+x^2}\right)} e^{-i(\Omega t - \varphi_0)}\Bigg\}, \end{aligned} \tag{20}$$

$$\begin{aligned} \nu = \frac{h_2}{\Delta H}\Bigg\{ & \frac{1 - \eta Q \frac{a}{2}\frac{y}{1+x^2}\left(\frac{1 - i(x-y)}{1 + x^2 - y^2 + 2iy\left(1 - \eta Q \frac{a}{2}\frac{x}{1+x^2}\right)}\right)}{x + y - i} e^{i(\Omega t - \varphi_0)} + \\ & + \frac{Q\, a\, y}{2(1+x^2)} \frac{1 - i(x+y)}{(x - y - i)\left(1 + x^2 - y^2 - 2iy\left(1 - \eta Q \frac{a}{2}\frac{x}{1+x^2}\right)\right)} e^{-i(\Omega t - \varphi_0)}\Bigg\}. \end{aligned}$$

The last expression of (20) enables us to determine directly the complex magnetic susceptibility of the specimen, χ_2, relative to a field of frequency $\omega = \Omega$. Indeed, substituting (20)

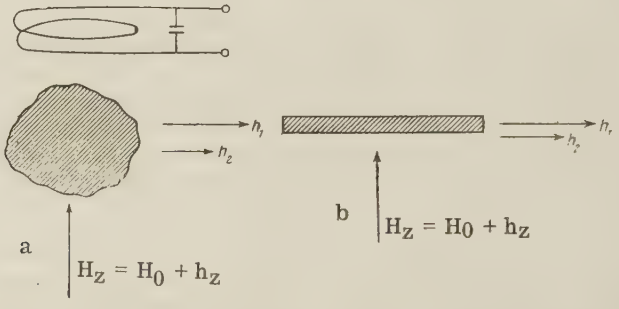


Figure 2. Two characteristic cases of coupling between a longitudinally polarized alternating field h_z and the corresponding component of the specimen magnetization m_z:
a) Field h_z is the field of the reaction of the resonant circuit interacting with the ferrite specimen; b) the field h_z is the demagnetizing field of the form of the specimen.

into (14) we can readily verify that χ_2 coincides with the coefficient of $\exp i(\Omega t - \varphi_0)$ in this expression, multiplied by $-M/h_2$, that is,

$$\chi_2 = -\frac{M}{\Delta H} \frac{1 - \eta Q \frac{a}{2}\frac{y}{1+x^2}\left(\frac{1 - i(x-y)}{1 + x^2 - y^2 + 2iy\left(1 - \eta Q \frac{a}{2}\frac{x}{1+x^2}\right)}\right)}{x + y - i}. \tag{21}$$

If

$$\chi_2 = \chi_2' - i\chi_2'', \tag{22}$$

then

$$\chi_2'' = \frac{M}{\Delta H (1+(x+y)^2)}\left\{1 - \eta Q \frac{a}{2}\frac{y}{1+x^2}\frac{1-(x^2-y^2)^2-4xy\left(1-\eta Q \frac{a}{2}\frac{x}{1+x^2}\right)}{(1+x^2-y^2)^2+4y^2\left(1-\eta Q \frac{a}{2}\frac{x}{1+x^2}\right)^2}\right\}. \tag{23}$$

For certain fixed relationships between x and y, and for a sufficiently large value of a, the imaginary part of the complex susceptibility χ''_2 can become negative, corresponding to amplification of the weak signal at the frequency $\omega = \Omega$. Indeed, let us put $x = 1$, $y = -2$ and $\eta a Q = 6$. Then $\chi_2'' = -1.75\frac{M}{\Delta H}$ The plot of the dependence of χ''_2 on y at $\eta a Q = 6$ is shown in Figure 3. This plot can be used to describe the frequency-band variation of this effect when the resonant circuit is being retuned in accordance with the variation of the frequency Ω.

Case 2. The field h_z is the demagnetizing field of the specimen shape. If the specimen is a thin plate magnetized in the short direction, then

$$h_z = -m_z. \tag{24}$$

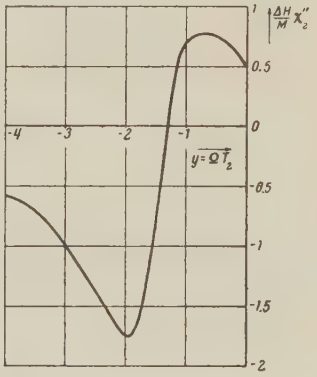


Figure 3. Dependence of the imaginary part of the complex magnetic susceptibility of the specimen, χ''_2, relative to a weak signal of frequency $\omega + \Omega$, on the quantity $y = \Omega T_2$ for the case when the field h_z is produced by a resonant circuit tuned to a frequency

Substituting (24) in (16) we get

$$\begin{aligned} T_2\dot{v} + (1+ix)v &= i\frac{a}{1+x^2}\Theta_1 + i\frac{h_2}{\Delta H}e^{i(\Omega t-\varphi_0)}, \\ T_2^2\ddot{\Theta}_1 + 2T_2\dot{\Theta}_1 + \left(1+x^2-a\frac{x}{1+x^2}\right)\Theta_1 &= \\ = \frac{h_2}{2\Delta H}\{(x-y+i)e^{i(\Omega t-\varphi_0)} + (x-y-i)e^{-i(\Omega t-\varphi_0)}\}, & \end{aligned} \tag{25}$$

whence,

$$\chi_2 = -\frac{M}{\Delta H}\frac{1-\frac{a}{2(1+x^2)}\frac{y-x-i}{1+x^2-y^2-a\frac{x}{1+x^2}+2iy}}{x+y-i}, \tag{26}$$

that is

$$\chi_2' = \frac{M}{\Delta H}\frac{\frac{a}{2(1+x^2)}\frac{1+x^2-y^2-a\frac{x}{1+x^2}-4xy+(y^2-x^2)(1+x^2-y^2)-x-y}{\left(1+x^2-a\frac{x}{1+x^2}-y^2\right)^2+4y^2}+(x+y)}{1+(x+y)^2}, \tag{27}$$

$$\chi_2'' = \frac{M}{\Delta H}\frac{1-\frac{a}{1+x^2}\left\{\frac{x\left(y^2-x^2-1+a\frac{x}{1+x^2}\right)-y(1+y^2-x^2)}{\left(1+x^2-y^2-a\frac{x}{1+x^2}\right)^2+4y^2}\right\}}{1+(x+y)^2}$$

When x = 0, the last expressions assume the simpler form

$$\chi_2' = -\frac{M}{\Delta H}\frac{y+\frac{y^2-1}{y^2+1}\frac{a}{2}}{1+y^2}, \qquad (28)$$

$$\chi_2'' = \frac{M}{\Delta H}\frac{1+\frac{y}{1+y^2}a}{1+y^2}.$$

When a = 6, x = 0, and y = -5 we have $\chi_2'' = -1.12\frac{M}{\Delta H}$.

The frequency dependence of χ'_2 and χ''_2 for a = 6 is shown in Figure 6. The same figure shows for comparison the frequency variations calculated for a = 0. A somewhat greater band variation can be obtained when x = 0.5.

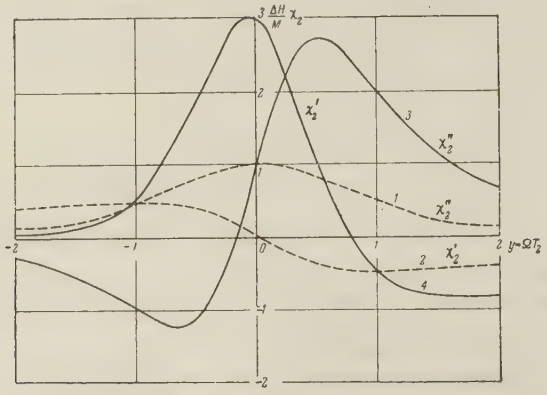


Figure 4. Dependence of the complex magnetic susceptibility of a specimen, $\chi_2 = \chi'_2 - i\chi''_2$, relative to a weak signal of frequency $\omega + \Omega$, on the quantity $y = \Omega T_2$ for the case when the field h_z is produced by the magnetizing action of the shape of the specimen:

1, 2) $H_0 = H_{res}$, a = 0; 3, 4) $H_0 = H_{res} - \Delta H$, a = 6.

5. Let us estimate tentatively the value of P_1, the microwave pumping power necessary to reach a value a = 6, and the weak-signal amplication coefficient obtained in this case. We assume that we use a specimen of yttrium ferrite measuring 3 • 3 • 0.3 mm, with saturation magnetization M = 1750 gauss and a resonance-line half-width ΔH = 1 oe. Then

$$P_1 = r_1 h_1^2 = a\frac{(\Delta H)^3}{M} r_1 = 3{,}43 \cdot 10^{-3} r_1. \qquad (29)$$

If the specimen is located in a waveguide one-quarter wavelength away from the short-circuited end, then the coupling coefficient r_1 between P_1 and H_1^2 at 3 cm wavelength can be set equal to 25 w/oe^2 [2]. In this case P_1 = 86 mw. The weak-signal amplification can be estimated by means of the coefficient

$$K = 1 + \frac{P_g}{P_2}, \qquad (30)$$

* The coefficient K characterizes the efficiency of the ferrite element in amplifying the weak microwave signal, since it defines the maximum power gain of the signal, obtainable at optimum matching of the load to the elements of the amplifier circuit. It must be distinguished from the current (or voltage) amplification coefficient at fixed load resistance.

where $P_g = -\mu_0 \frac{\chi_2'' h_2^2}{4} (\omega + \Omega) V$ denotes the power generated by the specimen at a frequency $\omega = \Omega$ under the influence of the stronger signal of frequency ω. Substituting in the formula the values $V = 2.7 \cdot 10^{-9}\ m^3$, $\mu_0 = 4\pi \cdot 10^{-7}$ henry/meter, we obtain $\omega + \Omega = 2\pi \cdot 10^{10}$, $h_2^2 = P_2/r_2$

$$K = 1 - 0.34 \frac{\chi_2''}{r_2}. \tag{31}$$

If $r_2 = r_1 = 25$ w/oe^2 and $\chi_2'' = -M/\Delta H = -1750$ then $K = 25$.

Thus, it becomes feasible to build, on the basis of the foregoing principle, a ferrite microwave power amplifier operating at pumping powers of less than 100 mw and providing a gain greater than 10 in a bandwidth for which the difference frequency between the signal and the local pumping generator Ω lies in the region of several megacycles.

The author is grateful to A. A. Pistol'kors who reviewed the manuscript and made several remarks on the substance of the investigation.

APPENDIX

Problem of Interaction Between Electromagnetic Signals and a Modulated Gyromagnetic Medium for Large Modulating-Field Amplitudes

We start from the Landau and Lifshitz equations in the form

$$\dot{m}_x + \frac{m_x}{T_2} = -\gamma (H_z m_y - h_y M_z), \tag{I}$$

$$\begin{aligned} \dot{m}_y + \frac{m_y}{T_2} &= \gamma (H_z m_x - h_x M_z), \\ \dot{m}_z + \frac{m_z}{T_1} &= \gamma (h_x m_y - h_y m_x). \end{aligned} \tag{II}$$

The notation is the same as in the text of the article. Introducing a new variable and making the substitution $N = Se^{\int \left(i\gamma H_z - \frac{1}{T_2}\right) dt}$ we reduce Equations (I) to the form

$$N = -i\gamma e^{\int \left(i\gamma H_z - \frac{1}{T_2}\right) dt} \int M_z (h_x + ih_y)\, e^{-\int \left(i\gamma H_z - \frac{1}{T_2}\right) dt} dt. \tag{III}$$

The quantity m_z is small compared with M and can be disregarded in (III). Substituting in (III) $M_z \simeq M$ and

$$h_x + ih_y = (h_1 + h_2 e^{i(\Omega t + \delta)})\, e^{i\omega t}, \quad h_z = h_0 \cos \Omega t \tag{IV}$$

and using the well-known formulas for series expansion [9]

$$e^{\pm iz \cos \Omega t} = \sum_{k=-\infty}^{\infty} (\pm 1)^k J_k(z)\, e^{ik\Omega t}, \tag{V}$$

we obtain

$$N = -i \frac{M}{\Delta H} e^{i\omega t} \sum_{p=-\infty}^{\infty} J_p(z)\, e^{ip\Omega t} \sum_{k=-\infty}^{\infty} (-1)^k \frac{J_k(z) h_1 - J_{k-1}(z) h_2 e^{i\delta}}{1 + iX_k} e^{ik\Omega t}. \tag{VI}$$

Here $X_k = x + ky$; $J_k(z)$ is the Bessel function of the first kind of order k whose argument $z = \frac{h_0}{\Delta H} \frac{1}{y}$. The initial phase angle δ is defined in formulas (IV) and (VI) relative to the field h_z.

Expression (VI) can be reduced by simple regrouping of the terms in the form

$$N = -i \frac{M}{\Delta H} \sum_{l=-\infty}^{\infty} A_l e^{i(\omega + l\Omega)t}, \tag{VII}$$

where

$$A_l = \sum_{k=-\infty}^{\infty} (-1)^l \frac{J_{k-l}}{1 + iX_k} \{J_k h_1 - J_{k-1} h_2 e^{i\delta}\}. \tag{VIII}$$

The relations obtained enable us to calculate the susceptibility of the medium χ_1 and χ_2 relative to signals at frequencies ω and $\omega + \Omega$, using the formulas

$$\chi_1 = -i\frac{M}{\Delta H}\frac{A_0}{h_1} = -i\frac{M}{\Delta H}\sum_{k=-\infty}^{\infty}\frac{J_k^2 - J_kJ_{k-1}\frac{h_2}{h_1}e^{i\delta}}{1+ix_k}, \tag{IX}$$

$$\chi_2 = -i\frac{M}{\Delta H}\frac{A_1}{h_2}e^{-i\delta} = -i\frac{M}{\Delta H}\sum_{k=-\infty}^{\infty}\frac{J_{k-1}^2 - J_kJ_{k-1}\frac{h_1}{h_2}e^{i\delta}}{1+iX_k}. \tag{X}$$

Rewriting (II) in the form

$$\dot{m}_z + \frac{m_z}{T_1} = \gamma\,\mathrm{Im}\,\{N(h_x - ih_y)\},$$

and using (VIII) and (VII), we obtain

$$m_z = -\frac{M}{(\Delta H)^2}\,\mathrm{Re}\sum_{l=0}^{\infty}\frac{B_l}{\frac{T_2}{T_1}+ily}e^{il\Omega t}, \tag{XI}$$

where

$$B_0 = A_0h_1 + A_1h_2e^{-i\delta}; \tag{XII}$$

$$B_l = (A_l + A_{-l}^*)h_1 + (A_{l+1}e^{-i\delta} + A_{1-l}^*e^{i\delta})h_2, \qquad l = 1, 2, \ldots.$$

The component m_Z at frequency Ω is described, in accordance with (XI) and (XII), by the expression

$$(m_z)_\Omega = -\frac{M}{\Delta H}\,\mathrm{Re}\sum_{k=-\infty}^{\infty}[h_1h_2(J_k^2e^{i\delta} + J_{k-1}J_{k+1}e^{-i\delta}) - (h_1^2J_kJ_{k+1} + h_2^2J_kJ_{k-2})]\frac{\xi}{(1-iX_k)(1+iX_{k+1})}. \tag{XIII}$$

The coefficient $\xi = \frac{2+iy}{\frac{T_2}{T_1}+iy}$ in (XIII) becomes equal to unity when $T_2 = 2T_1$ (condition for the conservation of the quantity M).

Formulas (IX), (X), and (XIII) are valid for practically all values of h_0, h_1, and h_2 (the only limitation is the condition $m_Z \ll M$).* They predict many new effects, which take place at large amplitudes of the modulating field h_Z, observed in [10]. These formulas, however, are somewhat too cumbersome for practical use.

If $h_0/2\Delta H \ll y$ then formulas (IX), (X), and (XIII) become simpler, and in the first approximation in the quantity $\frac{x}{2} = \frac{h_0}{2\Delta H}\frac{1}{y}$ assume the form

$$\chi_1 = -i\frac{M}{\Delta H}\frac{1}{1+ix}\left\{1 + \frac{h_2}{h_1}\frac{h_0e^{i\delta}}{2\Delta H}\frac{i}{1+i(x+y)}\right\}, \tag{XIV}$$

$$\chi_2 = -i\frac{M}{\Delta H}\frac{1}{1+i(x+y)}\left\{1 + \frac{h_1}{h_2}\frac{h_0e^{-i\delta}}{2\Delta H}\frac{i}{1+ix}\right\}, \tag{XV}$$

$$m_z = -\frac{M}{(\Delta H)^2}\,\mathrm{Re}\left\{\frac{h_1h_2e^{i\delta}}{(1-ix)[1+i(x+y)]} + \sum_{k=0}^{1}\frac{h_{k+1}^2h_0}{\Delta H}\frac{X_k}{1+X_k^2}\frac{1}{[1+i(X_k+y)][1-i(X_k-y)]}\right\}. \tag{XVI}$$

In the case when h_Z is the field of the reaction of the resonant circuit or of the specimen and is linearly related to m_Z, expression (XVI) can be used to determine the quantity $h_0e^{-i\delta}$. Substituting the expression thus obtained for $h_0e^{-i\delta}$ into formula (XV), we can obtain, for suitable values of the ratio of h_Z, expressions for χ_2, which coincide when $h_2 \ll h_1$ with the expressions derived in the text of the article (see formulas (21) and (27)).

---

* The analysis is carried out as applied to homogeneous precession. The possibility of occurrence of higher modes is disregarded.

REFERENCES

1. A. M. Mikaelyan, The problem of producing ferrite amplifiers for microwave frequencies. Radiotekhnika i Elektronika, 1958, 3, 11, 1323.
2. V. A. Fabrikov, Ferrite power amplifier for electromagnetic signals of high frequency. Radiotekhnika i Elektronika 1959, 4, 7, 1203.
3. V. A. Fabrikov, On the Effectiveness of Ferrite Elements as Microwave Mixers in Detector Circuits, Transactions of the Third All-Union Conference on Ferrites, Minsk, June 1959. "Ferriti" (Ferrites), Collection of articles, Academy of Sciences, Belorussian SSR Minsk, 1960, page 534.
4. L. Landau, E. Lifshitz, On the theory of the dispersion of magnetic permeability in ferromagnetic bodies, Phys. Z. Sowjetunion, 1935, 8, 153.
5. N. Bloembergen, Ferromagnetic Resonance in Nickel and Supermalloy, Collection of Translations, "Ferromagnetic Resonance, "edited by S. V. Vonsovskiy, IL, 1952.
6. G. V. Skrotskiy, V. T. Shmatov, On the thermodynamic theory of resonant and relaxation phenomena in ferromagnets, Zh. Eksp. Teoret. Fiz. 1958, 34, 3, 740.
7. V. A. Fabrikov, Possibility of amplifying the power of a weakly modulated signal by means of a gyromagnetic medium, Radiotekhnika i Elektronika, 1958, 3, 2, 190.
8. V. A. Fabrikov, On the conditions under which natural oscillations of magnetization arts in a gyromagnetic medium, Radiotekhnika i Elektronika, 1960, 5, 1, 117.
9. H. B. Dwight, Tables of Integrals and Other Mathematical Data, MacMillan, New York, 1947.
10. V. A. Fabrikov, V. T. Kadeyev, V. I. Kozlov, B. D. Kudryavtsev, Experimental Research of Effects Connected With Nutational Oscillations of Magnetization, Using Single Crystals of Yttrium Ferrite. Topics of Papers Delivered to the Conference on Ferromagnetism and Antiferromagnetism, Published by Academy of Sciences USSR, Leningrad, 1961.

Submitted to the Editors 6 December 1960

EQUIVALENT CIRCUITS OF THE DEFLECTING PLATES OF A CATHODE RAY TUBE

A. I. Sokolik *

It is shown that a single-mesh equivalent circuit of the deflecting-plate system is applicable only when the corresponding parameters LRC of the two plate systems are equal. An equivalent circuit is proposed, in which the deflecting plate systems are separated and in which the capacitances between the leads and the inductance and resistance of the circuit of the second anode is taken into account. The new equivalent circuit has been tested by simulation.

In oscillographic investigations of signals which are subject to sharp variations within a time on the order of 10^{-8}—10^{-9}sec, it is necessary to take into account the distortion due to the electric parameters of the circuit of the deflecting plates of the tube. These distortions as a rule exceed appreciably the distortions due to the finite transit time of the electrons in the field of the deflecting plates. Modern notions concerning the influence of the parameters of the system of plates on the waveform of the deflecting-plate voltage are discussed in [1] on the basis of the equivalent circuit of the deflecting system, shown in Figure 1, where R is the total resistance of the external circuit, including the internal resistance of the signal generator, L is the inductance of the plate leads and the signal leads, and C_0 is the capacitance of the deflecting plate.

---

* Deceased

If $R < 2\sqrt{L/C_0}$, and the applied signal is an ideal unit step of amplitude E, then the voltage getween the plates varies as

$$U_C = E\left[1 - e^{at}\left(\frac{a}{\omega}\sin\omega t + \cos\omega t\right)\right], \quad (1)$$

where

$$a = \frac{R}{2L}; \quad \omega = \sqrt{\frac{1}{LC_0} - \frac{R^2}{4L}}.$$

We shall define a resistance as optimal if it satisfies the condition

$$R_0 = \sqrt{\frac{2L}{C_0}} = \frac{0.22}{f_0 C_0}, \quad (2)$$

where

$$f_0 = \frac{1}{2\pi\sqrt{LC_0}}. \quad (3)$$

If the resistance is optimal, the time necessary for the voltage on the plates to build up from 0.1 to 0.9E is

$$t_b = 1.52\, R_0 C_0 \simeq \frac{0.35}{f_0}. \quad (4)$$

It is approximately 1.5 times smaller than in the case when the resistance is equal to the critical value. If the resistance is optimal and the leading edge of the input signal is ideally steep, an overshoot of 4 percent is produced. If the leading edge duration of the input signal is finite, the overshoot is smaller.

The internal resistance of the signal generator is usually smaller than optimal; therefore to avoid excessive overshoots one connects additional damping resistances in the plate circuit.

If we start from the equivalent circuit of Figure 1, it appears to us that the damping resistance can be connected, with identical results, into the circuit of any of the deflecting plates. Actually this is far from the truth. Figure 2 shows three oscillograms of a rectangular pulse of duration 10 nanoseconds, copied from the screen of a 13L05A oscilloscope tube under different connection conditions.

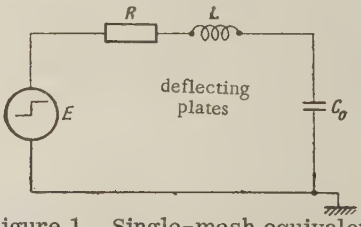
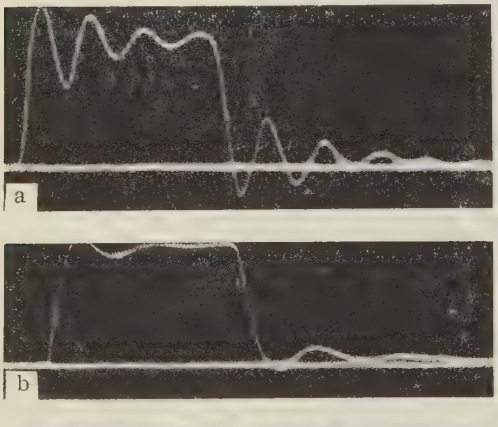


Figure 1. Single-mesh equivalent circuit for the deflecting plates.



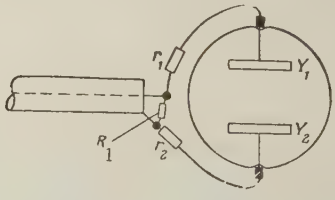
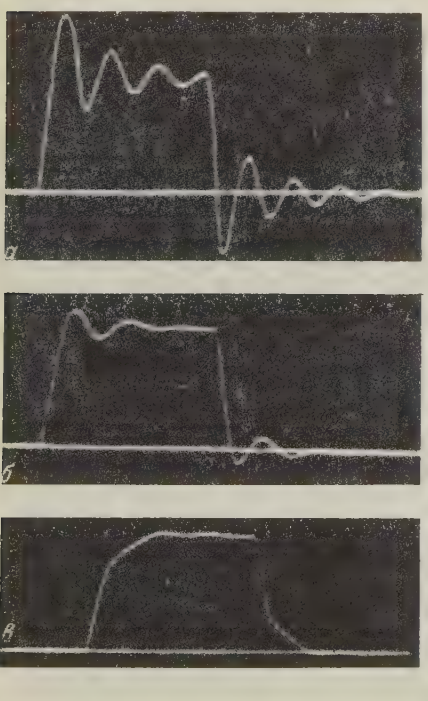


Figure 2. Oscillograms of rectangular pulse with duration 10^{-8} sec, compiled from the screen of a 13L05A tube: a) $r_1 = 0$, $r_2 = 0$, $f_2 = 350$ megacycles; b) $r_1 = 0$, $r_2 = 240$ ohms; c) $r_1 = 240$ ohms, $r_2 = 0$; d) circuit for applying the signal to the plate. RK-3 cable, $l = 14$ m, $R_L = 75$ ohms. The wires are 4.5 cm long each; all the resistors are of the ULM type.

The obvious influence that the location of the resistances exerts on the form of the signal leads to the conclusion that the equivalent network of the deflecting system must contain two RLC networks, corresponding to the circuit 5 of the two deflecting plates. Such a network, which we have proposed [2] is shown in Figure 3. The elements denoted by the subscript "1" pertain to the circuit of the "active" plate, while those designated "2" pertain to the circuit of the grounded plate; C_0 is the capacitance between plates; C_1 and C_2 the capacitances of the plates relative to the grounded electrodes of the tube (principally relative to the second anode and the screening partitions connected to it); C_l, the capacitance between the leads, and L_a and R_a are the inductance and resistance of the circuit connecting the second anode of the tube (both inside and outside the tube) to the external conductor of the cable over which the signal is fed, and which we assume to be the common point ("ground" on the equivalent network).

If the signal is applied to the deflecting plates symmetrically with respect to ground, then the EMF of the signal generator in the circuit of the first plate in the network of Figure 3 should be taken equal to $E_1 = E/2$, but in the circuit of the second plate there should be connected a second generator with EMF $E_2 = -E_1$.

A mathematical analysis of the transient characteristics of the proposed equivalent networks is in general difficult to carry out, and therefore these circuits were investigated by simulation. The capacitances C_0, C_1 and C_2 of the model were determined by measuring the respective capacitances of the lower deflecting plates D_3 and D_4 of the 13L05A tube; the inductances L_1 and L_2 were determined from the natural frequencies of the deflecting-plate circuits. The capacitances and inductances were increased 1000 times the actual values, thereby increasing the time scale by the same factor. Figure 4 shows oscillograms obtained with a model of the deflecting system, constructed in accordance with the circuit of Figure 3, for different connections of the damping resistance, corresponding to conditions under which the oscillograms shown in Figure 2 were obtained. In the simulation account was taken of the resistance of the deflecting-plate circuit, due to surface effects and due to losses to radiation. It was calculated from the oscillograms of the natural oscillations of the deflecting system of the tube at low values of total resistance ($R_1 = R_2 = 13.5$ ohm), and its value was approximately 7 ohms. Comparison of the oscillograms taken from the screen of the tube and those obtained with the model is evidence that the equivalent network of Figure 3 agrees well with the real circuit of the deflecting plates of the tube.



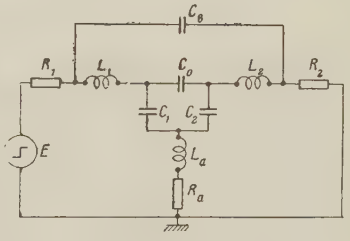


Figure 3. Two-mesh equivalent circuit of the deflecting-plate circuits.

Figure 4. Oscillograms obtained with a model of the deflecting-plate circuits using the equivalent network of Figure 3, with $C_0 = 1500$ $\mu\mu$f, $C_1 = C_2 = 2100$ $\mu\mu$f, $L_1 = L_2 = 43$ μh, $C_l = 500$ $\mu\mu$f, $L_a = 45$ μh, $R_a = 50$ ohms. Input pulse 10 msec. Impedances in the plate circuits: a) $R_1 = 44$ ohms, $R_2 = 7$ ohms; b) $R_1 = 44$ ohms, $R_2 = 247$ ohms; c) $R_1 = 284$ ohms, $R_2 = 7$ ohms.

A model constructed in accordance with the equivalent circuit of Figure 3 but without the elements C_l, L_a and R_a also duplicates well the waveform of the voltage on the deflecting plates, with the exception of the case when the model included a damping resistance of 240 ohms only in the circuit of the first plate. The oscillogram obtained in the circuit of the first plate. The oscillogram obtained in that case (Figure 5) differs appreciably from the oscillogram of Figure 2c; this indicates that in some cases the capacitance between the leads of the deflecting plates of the tube and the impedance of the circuit of the second anode appreciably influence the waveform of the voltage on the deflecting plates.

Let us discuss the case when all the parameters of the circuits of the two plates are equal. When $L_1 = L_2 = L$, $R_1 = R_2 = R$, $C_1 = C_2 = C$ the transform of the voltage on capacitance C_0 in response to a unit step E applied to the input of the circuit of Figure 3 has the following form:

$$U_{C_0}(p) = \frac{E}{p^3 2LC_e C_l R + p^2 LC_e + pR(C_e + C_l) + 1}, \tag{5}$$

where

$$C_e = C + 2C_0. \tag{6}$$

It follows from (5) that for the case under consideration the voltage on the plates is independent of the parameters of the second-anode circuit. This gives us the right to assume $R_a = \infty$, which is equivalent to replacing the circuit of Figure 3 by a much simpler circuit, shown in Figure 6. If we put at the same time $C_l = 0$, we arrive at the single-mesh circuit (Figure 1), comprising the elements 2L, $_2$R, and $C_e/2$, or what is the same, of the elements L, R, and C_e.

Thus, the single-mesh equivalent circuit of the deflecting system, shown in Figure 1, can be used provided the resistances, inductances, and capacitances in the circuits of the two plates are equal, and provided the capacitance between the leads can be neglected. Under these conditions expressions (1) — (4) are valid if one assigns the circuit parameters the values indicated above. In the case when the parameters of both plates are equal (with $C_l = 0$ and $Z_a = 0$), the optimal value of the inductances necessary to obtain a minimum buildup time for the voltages on the plates, with a permissible overshoot $\delta = 4$ percent, will be

$$L_{1\,\mathrm{opt}} = L_{2\,\mathrm{opt}} = \frac{C_e(R_i + r_{01})}{2}, \tag{7}$$

where R_i is the internal resistance of the generator; r_{01} is the resistance of the first plate circuit, due to the resistance of the lead in and to radiation. If the inductance is less than

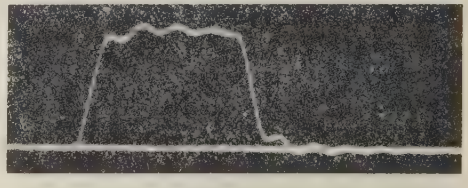


Figure 5. Oscillogram obtained with a model under the same conditions as for Figure 4c $C_l = 0$, $L_a = 0$, and $R_a = 0$.

optimal, the overshoot decreases, but the buildup time of the voltage increases. In the limit, when $L_1 = L_2 = 0$, we have

$$t_b = 0.22(R_i + r_{01})C_e, \tag{8}$$

which is 1.45 times greater than when condition (7) is satisfied. However the case when the parameters of the circuits of the two plates are equal is not optimal. Tests with a model

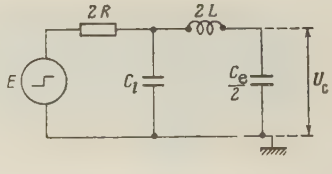


Figure 6. Equivalent network of the deflecting-plate circuit with $L_1=L_2=L$, $R_1=R_2=R$, $C_1=C_2=C$

based on the equivalent circuit of Figure 3 have shown that the resistance and inductance in the circuit of the first (active) plate influence the buildup time of the voltage and the overshoot to a greater degree than the resistance and inductance in the circuit of the grounded plate (Figures 7 and 8), it is therefore advantageous to decrease the first inductance even at the cost of increasing L_2. In this case the optimal value of the resistance in the circuit of the first plate decreases, and this provides an appreciable gain in the voltage buildup time and in many cases eliminates the need for a special damping resistance in the circuit of the first plate. The damping resistance in the circuit of the second plate is as a rule essential.

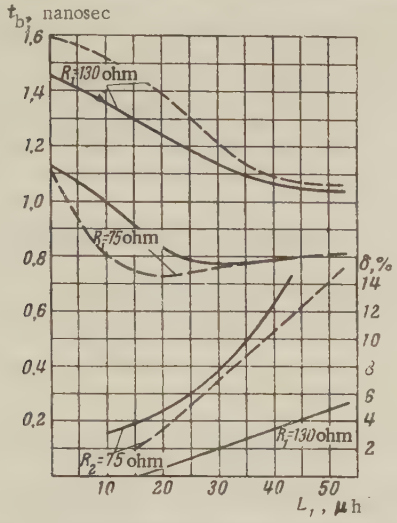


Figure 7. Plot showing the dependence of the buildup time t_b of the voltage and the overshoot δ on the inductance L_1 with L_2 = const (L_2 = = 43 μh, R_2 = 130 ohms, C_0 = 1.5 μμf, $C_1 = C_2$ = = 2.1 μμf):

Solid curves — C_l = 0, L_a = 0, r_a = 0; dashed curves — C_l = 0.5 pf, L_a = 45 μh, r_a = =50 ohms.

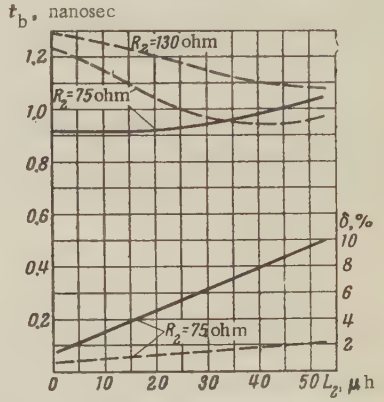


Figure 8. Plot showing the dependence of the voltage buildup time t_b and of the overshoot δ on the inductance L_2 with L_1 = const (L_1 = 43 μh, R_1 = 130 ohms, C_0 = 1.5 μμf, $C_1 = C_2$ = 2.1 μμf):

Solid curves — C_l = 0, L_a = 0, r_a = 0; dashed curves — C_l = 0.5 μμf, L_a = 45 μh, r_a = 50 ohms.

The effect of the capacitance between leads is illustrated by the oscillograms of Figure 9, obtained with the model at equal inductances and at the optimal value of the resistance. The oscillogram of Figure 9a is obtained with C_l=0, while values C_l=0.5 and 1.5 micromicrofarads

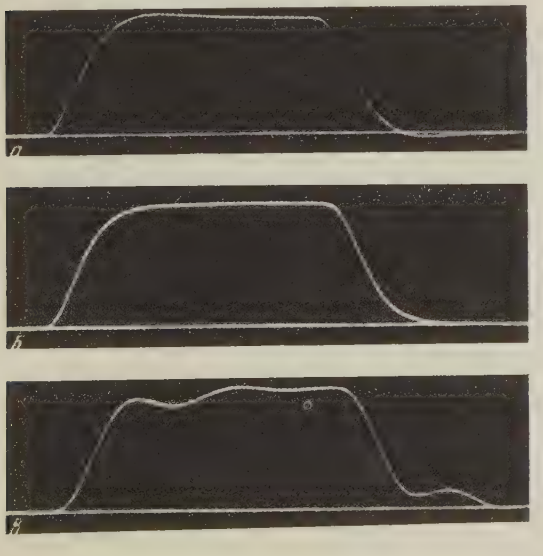


Figure 9. Oscillograms obtained with the model at L_1=L_1=43 μh, C_0=1500 μμf, C_1= C_2=2100 μμf, R_1=R_2=130 ohms, input pulse 5 microsecond:

a) C_l=0. b) C_l=500 μμf; c) C_l=1500 μμf.

Figure 10. Rectangular pulse of duration $1.5 \cdot 10^{-8}$ sec, copied from the screen of a 13LO5A tube.

were used in the succeeding oscillograms The duration of the input pulse when recalculated to the time scale of the tube is 5 nanoseconds. With C_l=0.5 micromicrofarad (that is, when C_l=0.1 C_e) the overshoot decreases from 4 percent to 0 percent to 0 practically without an increase in the voltage buildup time, but further increase in the capacitance greatly distorts the start of the flat top of the pulse, and a negative overshoot, as it were, appears. The same effect is producted by the capacitances of the leads with respect to ground.

A study of the influence of individual elements of the deflecting-plate circuits on the waveform of the signal, carried out by simulation on the basis of the equivalent circuit proposed, has made it possible to apply the signal to the tube 13L05A (in which the leads to the plates are in opposite sides of the neck) in such a way, that the buildup time of the pulse leading edge was 1.1 nanosecond on the oscillogram (see Figure 10). The same pulse, recorded with a tube having a traveling wave deflecting system with a bandwidth of 500 Mc, had a voltage buildup time of approximately 0.8 namosecond.

REFERENCES

1. I. Lewis and F. Wells, Millimicrosecond Pulse Techniques (Russian Translation) IL, 1956, pages 205 — 217.
2. A.I. Sokolik, Pulse Techniques — A Textbook, MIFI Press, pages 202 - 206.

Submitted to the Editors 15 January 1960

ELECTROSTATIC QUADRUPOLE LENS, THE FIELD OF WHICH HAS TWO ANTISYMMETRY PLANES AND NO SYMMETRY PLANE

A. M. Strashkevich

The distribution of the potential and the electron-optical properties of an electrostatic quadrupolse lens, which has no symmetry planes but does have antisymmetry, are considered.

In [1] is described a magnetic quadrupole lens having two antisymmetry planes but no symmetry planes, and its electron-optical properties are investigated. We consider an electrostatic quadrupole lens having the same kind of antisymmetry. For this purpose the azimuthal symmetry planes of each pair of electrodes should make not a right angle with each other, as usual, but a different angle 2γ (Figure 1). Then the field has no symmetry planes, but has two mutually perpendicular antisymmetry planes (in Figure 1 these are the coordinate planes xz and yz). Both when $2\gamma = \pi/2$ and for arbitrary γ in each azimuthal plane (that is, in each plane passing through the z-axis) the potentials of the points that are symmetrical with respect to the z-axis are of equal magnitude and sign.

Denoting the constant value of the potential on the z-axis (and meaning also in all the points of the antisymmetry planes xz and yz) by Φ_0, we can represent the potential at any point of the field under consideration by means of the expression

$$\Phi(x, y, z) = \Phi_0 + f_{11}(z)\,xy + f_{13}(z)\,xy^3 + f_{31}(z)\,x^3y + \cdots . \tag{1}$$

Let us change over the coordinates x, y, z to the coordinates ξ, η, z wherein the ξ and η axis are turned relative to the x- and y-axis by an angle $\pi/4$ (see Figure 1); the coordinates are interrelated by the equations

$$x = \frac{\sqrt{2}}{2}(\xi - \eta), \quad y = \frac{\sqrt{2}}{2}(\xi + \eta), \tag{2}$$

$$\xi = \frac{\sqrt{2}}{2}(x + y), \quad \eta = \frac{\sqrt{2}}{2}(-x + y). \tag{2a}$$

Then (1) becomes

$$\Phi(\xi, \eta, z) = \Phi_0 + \frac{1}{2} f_{11}(z)(\xi^2 - \eta^2) + \frac{1}{4}[f_{13}(z) + f_{31}(z)](\xi^4 - \eta^4) + \frac{1}{2}[f_{13}(z) - f_{31}(z)](\xi^3\eta - \xi\eta^3) + \cdots. \quad (3)$$

It is seen from this expression that near the z-axis (where the powers of ξ and η above the second can be neglected), the planes ξz and ηz are symmetry planes (since the coordinates enter only in even powers), but at considerable distances from the axis (that is, when terms of fourth order cannot be neglected) this symmetry is lacking, provided the coefficients f_{13} and f_{31} are not identical with each other; the latter takes place when $\gamma = \pi/4$, and then the field is symmetrical over all of space relative to the planes ξz and ηz. Thus, if for accidental reasons (for example, owing to inaccuracy in installation) the angle γ deviates from $\pi/4$, the electron-optical properties of the central region of the field of the system will not change qualitatively, for in this region neither the double antisymmetry nor the double symmetry of the field are violated. The quantitative changes in the electron-optical parameters, which take place in this case, will be considered below.

Let us illustrate the foregoing general premises with a specific example of a system of four spherical electrodes of small radius (Figure 1). If we assume the spheres to be point-

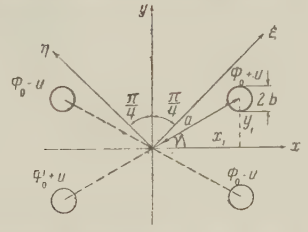


Figure 1. Example of a quadrupole lens having two planes of antisymmetry (xz, yz) in the field but no symmetry planes.

like and denote their charges by $\pm q$, the potential at the point with coordinates x, y, z, is

$$\Phi = \Phi_0 + q\{[(x - x_1)^2 + (y - y_1)^2 + z^2]^{-\frac{1}{2}} + [(x + x_1)^2 + (y + y_1)^2 + z^2]^{-\frac{1}{2}} - [(x - x_1)^2 + (y + y_1)^2 + z^2]^{-\frac{1}{2}} - [(x + x_1)^2 + (y - y_1)^2 + z^2]^{-\frac{1}{2}}\}. \quad (4)$$

Expanding (4) in a Maclaurin series and comparing with (1), we obtain

$$f_{11}(z) = \left(\frac{\partial^2\Phi}{\partial x\,\partial y}\right)_{00} = 12qx_1y_1(x_1^2 + y_1^2 + z^2)^{-\frac{5}{2}} = 6qa^2\sin 2\gamma\,(a^2 + z^2)^{-\frac{1}{2}}, \quad (5)$$

$$f_{13}(z) = \frac{1}{6}\left(\frac{\partial^4\Phi}{\partial x\,\partial y^3}\right)_{00} = 5qa^2\sin 2\gamma\,(a^2 + z^2)^{-\frac{9}{2}}[a^2(7\sin^2\gamma - 3) - 3z^2], \quad (5a)$$

$$f_{31}(z) = \frac{1}{6}\left(\frac{\partial^4\Phi}{\partial x^3\,\partial y}\right)_{00} = 5qa^2\sin 2\gamma\,(a^2 + z^2)^{-\frac{9}{2}}[a^2(7\cos^2\gamma - 3) - 3z^2]. \quad (5b)$$

Comparing (5a) and (5b) we find that when $\gamma = \pi/4$ these expressions coincide; consequently, terms with odd powers of ξ and η vanish in formula (3), that is, the field will have symmetry with respect to the planes ξz and ηz in regions not restricted to the near-axial (central) region. Expression (5) coincides when $\gamma = \pi/4$ with that given in [2]. Here, we replace q by the expression bU, that is, we assume that the potential on the surface of a sphere of small radius b is due only to the point charge located at the center of the sphere, so that this surface is equipotential and can be replaced by the surface of a metallic electrode with potential q/b (see [2-4]).

The equations of the trajectory of a relativistic charged particle [3] [formulas (1.2a)] for the central region can be written in the following form:

$$x'' = k(z)\,y, \quad y'' = k(z)\,x; \quad (6)$$

the primes denote differentiation with respect to z, and the variable coefficient k(z) has a value

$$k(z) = \frac{1 - \frac{e\Phi_0}{m_0c^2}}{1 - \frac{e\Phi_0}{2m_0c^2}}\frac{f_{11}(z)}{2\Phi_0} = A\frac{f_{11}(z)}{2\Phi_0}, \quad (6a)$$

where e is the algebraic value of the charge; m_0 is the rest mass of the particle; c is the velocity of light in vacuum. To integrate Equations (6) we can use a successive-approximation method similar to that used in [3] (Section 15) for the calculation of trajectories in fields with four symmetry planes and four antisymmetry planes. Putting $x' = X$ and $y' = Y$, and letting x_0 and y_0 stand for the values of x and y when $z = z_0$, we get

$$x(z) = x_0 + \int_{z_0}^{z} X(z)\,dz, \qquad y(z) = y_0 + \int_{z_0}^{z} Y(z)\,dz. \tag{7}$$

But as a result of (6) we have

$$X(z) = x_0' + \int_{z_0}^{z} k(z)\,y\,dz, \qquad Y(z) = y_0' + \int_{z_0}^{z} k(z)\,x\,dz. \tag{8}$$

Taking as the first approximation

$$x_1 = x_0 + x_0'(z - z_0), \quad y_1 = y_0 + y_0'(z - z_0)$$

(i.e., $X_1 = x_0'$, $Y_1 = y_0'$)

and substituting in (8) we obtain X_2 and Y_2, which we substitute in (7) in order to determine the coordinates x_2, Y_2 of the trajectory point in the second approximation. Continuing this process we obtain the third, fourth, etc. approximation. For the specific system considered above (measuring z from the center of the quadrupole lens and taking $z_0 = -l$), the second approximation has the form

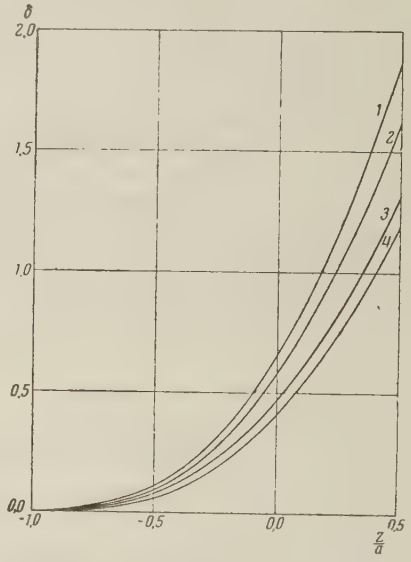


Figure 2. Projections on the symmetry plane of the central part of the field of the near-axial trajectories, initially parallel to the z-axis, at different values of γ:
1 — $\gamma = \pi/4$; 2 — $\gamma = \pi/6$; 3 — $\gamma = \pi/8$; 4 — $\gamma = \pi/9$

$$x_2 = x_0 + x_0'(z + l) + A\frac{U}{\Phi_0} b \sin 2\gamma\left(Z_1\frac{y_0}{a} + Z_2 y_0'\right), \tag{9}$$

$$y_2 = y_0 + y_0'(z + l) + A\frac{U}{\Phi_0} b \sin 2\gamma\left(Z_1\frac{x_0}{a} + Z_2 x_0'\right), \tag{9a}$$

where

$$Z_1 = \frac{1 + 2\frac{z^2}{a^2}}{\sqrt{1 + \frac{z^2}{a^2}}} - \frac{1 - 3\frac{l}{a}\frac{z}{a}\left(1 + \frac{2}{3}\frac{l^2}{a^2}\right)}{\sqrt{\left(1 + \frac{l^2}{a^2}\right)}};$$

$$Z_2 = \frac{\frac{l}{a}\left(1 + 2\frac{z^2}{a^2}\right) - \frac{z}{a}}{\sqrt{1 + \frac{z^2}{a^2}}} + \frac{\frac{z}{a}\left(1 + 2\frac{l^2}{a^2}\right) - \frac{l}{a}}{\sqrt{1 + \frac{l^2}{a^2}}}.$$

If we use the coordinates ξ and η, then the variables in the equations for the near-axial trajectory separate:

$$\xi_2 = \xi_0 + \xi_0'(z + l) + A\frac{U}{\Phi_0} b \sin 2\gamma\left(Z_1\frac{\xi_0}{a} + Z_2\xi_0'\right), \tag{9b}$$

$$\eta_2 = \eta_0 + \eta_0'(z + l) - A\frac{U}{\Phi_0} b \sin 2\gamma\left(Z_1\frac{\eta_0}{a} + Z_2\eta_0'\right). \tag{9c}$$

Figure 2 shows the z dependence of the quantity δ, which stands for the expression $\frac{\xi_2 - \xi_0}{\xi_0} : A\frac{U}{\Phi_0}\frac{b}{a}$ (or $\frac{\eta_0 - \eta_2}{\eta_0} : A\frac{U}{\Phi_0}\frac{b}{a}$); the curves obtained represent the projections of the near-axial trajectories on the symmetry plane for $\xi_0' = \eta_0' = 0$ and $l = a$ at different values of γ.

As is well known, the electron-optical parameters of an ordinary quadrupole lens, the field of which has two symmetry planes and two antisymmetry planes, can be expressed in the Gaussian approximation (that is, for the near-axial rays) in terms of the coefficient of the difference $\xi^2 - \eta^2$. in (3). In [3,4] this variable coefficient is denoted by C_z; for a system of four small spheres it has a value $3a^2bU(a^2+z^2)^{-\frac{5}{2}}$, which we obtained from the formulas of the present paper by taking $\gamma = \pi/4$, since it follows from (3) and (5) that

$$C_z = \frac{1}{2} f_{11}(z) = 3a^2bU \sin 2\gamma (a^2 + z^2)^{-\frac{5}{2}}. \quad (10)$$

From this expression and from [3] [formulas (15.47) and (15.48)] we see that for any value of γ the coordinates of the principal points are unchanged, and the absolute magnitude of the optical strengths changes in proportion to sin 2γ.

Conclusions. The central region of a doubly antisymmetrical field has two symmetry planes, even if there exist no field symmetry planes in a wider region.

The equations of the trajectory ot a relativistic charged particle in the near-axial region can be represented in the form (6); it is convenient to solve these equations by the method of successive approximations.

For a specific system comprising four small charged spheres, we have found the coefficients needed for the calculation of the electron-optical properties, and obtained the trajectories in the second approximation.

For arbitrary variation of the angle 2γ between the azimuthal planes of symmetry of the system of spherical electrodes, the coordinates of the principal points are invariant, and the optical strengths change in proportion in sin 2γ.

REFERENCES

1. M.G. Markovich, I.I. Tsukkerman, Zh. Tekh. Fiz. 1960, 30, 11, 1362.
2. G.D. Archard, Proc. Phys. Soc. B, 1955, 68, 11, 817.
3. A.M. Strashkevich, Electron Optics of Electrostatic Fields Having No Axial Symmetry, GIFML, 1959.
4. A.M. Strashkevich, Radiotekhnika i Elektronika, 1960, 5, 12, 1997.

Submitted to the Editors 28 January 1961

INVESTIGATION OF SOME PROPERTIES OF COLD MAGNESIUM OXIDE CATHODES WITH SELF MAINTAINING (MOLTER-TYPE) EMISSION

N.Ya. Basalayeva, T.M. Yekimenko, M.I. Yelinson, D.V. Zernov, Ya.V. Savitskaya, A.A. Yasnopol'skaya

A technology is developed for the production of layers of magnesium oxide, and a study is made of the effect of various factors on the self-maintaining emission of the cathode. The effect of thickness of coating on the emitting properties of the cathode is determined. It is shown that the best emitting ability is reached at an optimum amount of absorbed oxygen. It is determined that the times of formation of emission is investigated. A microscopic investigation is made of the distribution of the emission over the surface of the cathode. The effect of a magnetic field perpendicular to the surface of the layer of magnesium oxide on the emission is established. The mechanism of the self-maintaining emission is discussed.

INTRODUCTION

Recently several communications have been published [1 — 5] on the development and successful use of a new type of cathodes for vacuum tubes; in this cathode the emission mechanism is connected with processes which take place in dielectric layers in the presence of a strong electric field. The main advantage of this cathode compared with thermionic cathodes is that it operates in a cold state and needs no heating energy, and that the emission is established within a relatively short time after turning on the voltage. The emission results from brief illumination of the cathode or brief bombardment of its surface with fast primary electrons from an auxiliary source (starter) within a time on the order of several tenths of a second, after which self-maintained emission (Molter-type emission) is established (the emission does not cease after cessation of irradiation with light or with electrons).

It must be noted that the phenomenon of self-maintained emission from thin dielectric films has been known for a relatively long time. Thus, one of the authors of the present article (D. V. Zernov) observed as long ago as 1937 [6] that when thin transparent films of B_2O_3 deposited on a metallic (nickel) base are bombarded with primary electrons a secondary current is produced, which exceeds the primary current by hundreds of times. This current is maintained for an indefinitely long time after the primary electron beam is turned off, and unlike the alternating emission from films of Al_2O_3 observed earlier by Molter, the emission observed by us showed no tendency to decrease. Later on the phenomenon of self-maintaining emission was observed in other dielectric layers, including MgO [7,8], but for a long time this effect found no practical application, owing to the insufficient stability and reproducibility. Only in recent years, as a result of searches for an optimum technology for preparation of cathodes with self-maintaining emission, were conditions found under which it is possible to produce cathodes in the form of a porous layer of magnesium oxide on a magnesium oxide on a metallic (nickel) core, the quality of which (service life, reproducibility, stability, uniformity of distribution of the emission over the surface) turned out to be sufficiently high to be able to make practical the use of the new cathode in vacuum tubes.

The literature describes the experience in employing cold magnesium-oxide cathodes in vacuum tubes, procedures for its preparation, its structure, and its basic properties. Volt-ampere characteristics are given for the emission, data are cited on the average cathode-surface potential, on the distribution of the emitted electrons by energies, etc.

The purpose of the present investigation was a study of certain properties of cold magnesium oxide cathodes, which have not been sufficiently elucidated in the hitherto published papers.

Among the problems to be solved were: 1) determination of the role of treating the MgO layer in an oxygen atmosphere; 2) investigation of the temperature dependence of the emission; 3) study of the process of formation of the emission when the anode voltage is turned on; 4) determination of the effect of a magnetic field on the emission; 5) determination of the connection between the emitting ability and the thickness of the magnesium oxide layers; 6) study of the character of the distribution of emission over the surface of the cathode; 7) determination of the possibility of exciting self-maintained emission by direct application of voltage on the MgO layer; 8) determination of the effect of the trapping field on the emission.

On the basis of the results of these investigations, we advance below certain notions concerning the mechanism of self-maintaining emission and the prospect of its practical utilization.

1. EXPERIMENTAL PROCEDURE AND RESULTS

The cathodes were made by two methods, cathophoresis and spraying.

A study of the stability and scattering ability of suspensions of magnesium carbonate of various compositions, and a determination of the connection between the composition and the quality of the coating obtained in different modes has enabled us to recommend for depositing coatings by the cathophoresis method a 5 percent suspension of normal magnesium carbonate in absolute methanol and in a mixture of methanol with acetone in 1:1 and 3:1 ratios.

For spraying we used an amyl-acetate suspension of a mixture of basic magnesium carbonate and magnesium peroxide.

The cores used were electrolytic nickel, nickel with additives (magnesium, etc.), platinum-coated nickel and platinum. The results given below pertain essentially to cathode with cores made of brand NM nickel (with magnesium added) and platinum-coated nickel. The thickness of the layers varied over a wide range, from 12 to 16 microns for sprayed coatings and from 6 to 35 microns for cathophoretic ones. The shrinkage of the layers after calcination in vacuum was 20-30 percent. Layers obtained either by spraying or by cathophoresis had relatively high porosity. According to an approximate estimate, made on the basis of the ratio of the density of the resultant layers to the density of the corresponding single crystals, the total volume of pores in the layer (prior to vacuum treatment) was about 80 percent of the total volume of the layer. In spite of the shrinkage of the layers after heat-treatment, the percentage of pores did not decrease in this case, and even increased somewhat. The roughness of the layers was about 3 microns for cathophoretic cathodes and about 10 microns for sprayed ones. The most probable grain dimension prior to heat-treatment of the cathode is estimated to be about 1 micron. The structures of the sprayed and cathophoretic cathodes differ somewhat from each other (Figures 1 and 2).

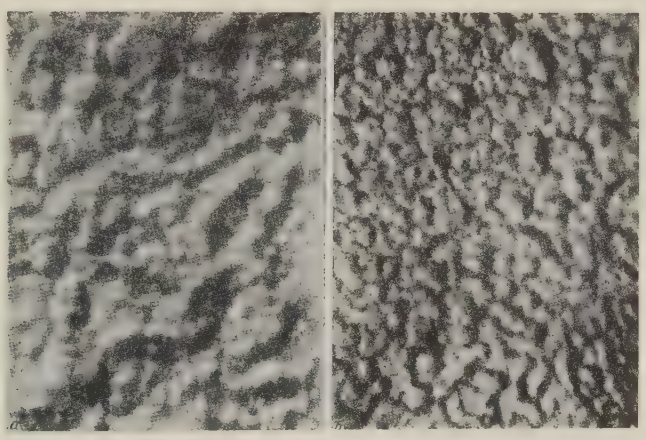


Figure 1. Photomicrographs of the surface of a sprayed cathode: a) prior to heating in vacuum; b) after heating in vacuum (magnification 115 ×).

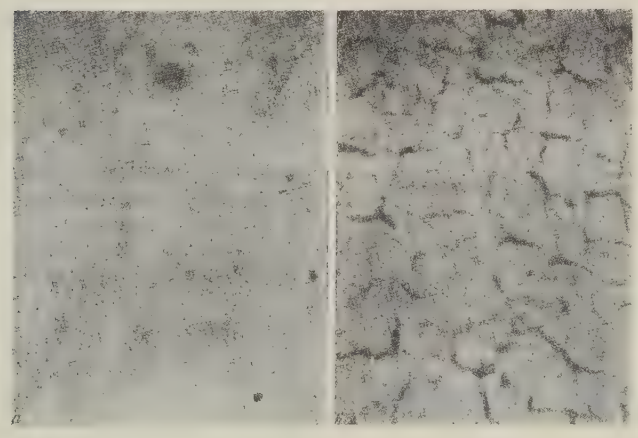


Figure 2. Photomicrograph of the surface of a cathophoretic cathode: a) prior to heating in vacuum; b) after heating in vacuum (magnification 115×).

The experimental tubes were for the most part prepared in the form of a simple diode with a tubular cathode of oval cross section and with a mesh anode. The emission was initiated in these diodes by means of a starter in the form of a thin tungsten wire (100 microns in diameter) heated by means of current. Some experiments were carried out in tubes in which the emission was initiated by bombarding with electrons from a special electron gun.

The treatment of the cathodes consisted of heating the layer at 850°C in vacuum for 10 min. This brought about decomposition of the magnesium carbonate and formation of the magnesium oxide layer. This was followed by an essential operation, namely heating in an oxygen atmosphere.

The experiments were carried out with a large number of cathode specimens (about four hundred). The data obtained on the current density (up to 50 ma/cm^2 for a short time and about 5 ma/cm^2 steadily with a total current 10 — 100 ma), on the degree of homogeneity of the distribution of the emission over the surface of the cathode and on the stability of emission with time, all gave evidence that we were dealing with cathode specimens with parameters already known in the literature.

By way of example, Figure 3 shows the volt-ampere emission characteristics, plotted for different cathodes.

a. Effect of oxygen. Experience has shown that layers not treated in oxygen did as a rule not give rise to emission, and whatever emission did take place occurred only in a few individual centers and was quite unstable. Yet it is not clear whether this oxygen treatment is necessary for oxidizing the impurities contained in the layer so as to impart to it a high dielectric strength (a very unlikely cause), for oxidizing the nickel under the layer of magnesium oxide, or whether the presence of axygen in the layer is for one reason or another an important factor in itself.

To answer these questions we undertook the following experiments.

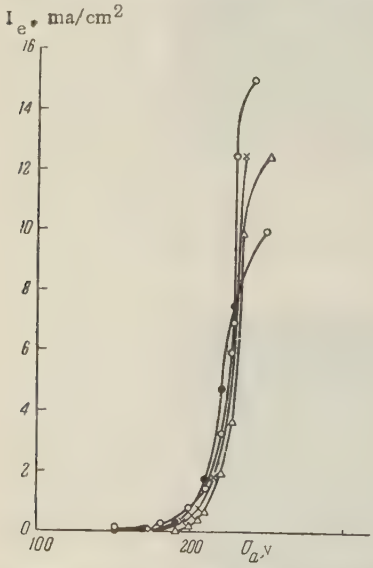


Figure 3. Volt-ampere characteristics of different cathodes.

After vacuum heat-treatment, the cathodes were subjected to several cycles of treatment in oxygen, each cycle consisting of heating the cathodes in an oxygen atmosphere at an initial pressure 0.1 mm Hg (850°C), continuing until the oxygen pressure in the apparatus dropped to 10^{-3} mm Hg, which took about three minutes (control experiments have shown that the bulk of the oxygen is absorbed in this case not by the magnesium oxide layer, but by the core and by the various parts of the tube). After each cycle of processing in oxygen, we measured the maximum cathode emission current I_e, above which the emission became unstable, and the voltage drop U_a, corresponding to this current, between the cathode and anode of the diode. The ratio I_e/U_a was adopted as some criterion characterizing the quality of the cathode. The choice of this criterion is naturally somewhat arbitrary (inasmuch as the connection between I_e/U_a has in this case a nonlinear character and consequently the ratio I_e/U_a is not an analog (in the full sense of the word) of the conductivity at the cathode saturation current); however, for lack of anything better it does give a correct idea of the quality of the cathode, which obviously will be the better, the greater the emission current that the cathode can produce and the less the anode voltage at which this current can be obtained.

The results obtained are illustrated by plots of I_e/U_a vs. N — the number of processing cycles. These plots are shown in Figure 4 (for cathophoretic coating) and Figure 5 (for sprayed coating). The character of the curves is in general the same on both plots. Regardless of the thickness and method of depositing the layer, relatively rapid growth in the ratio I_e/U_a is observed with an increasing amount of absorbed oxygen; at a certain degree of saturation with

oxygen, the curves passed through a maximum, followed by a rather rapid decrease. This is evidence first of all not only of the appreciable significance of the operation of treating the cathode with oxygen, but also of the high sensitivity of the cathode to this operation. The curves shown in Figure 4 disclose a regular variation of the position and magnitude of the maximum with increasing thickness of the film. In Figure 5 this regularity is less pronounced.

In order to exclude the assumption made above, namely that the role of processing with oxygen reduces to the formation of an intermediate NiO layer, we tested cathodes made with coares of platinum-coated nickel (thickness of platinum coating 50 microns). However, the character of the dependence of I_e/U_a on the number of processing cycles remained the same for this cathode (see Figure 6). That the role of the oxygen cannot be reduced to the oxidation of the nickel or some other impurities in the layer itself is also evidenced by the results of an experiment consisting of alternating cycles of heating in vacuum and in oxygen, illustrated by the plot shown in Figure 7. This figure shows that the cathode treated in oxygen

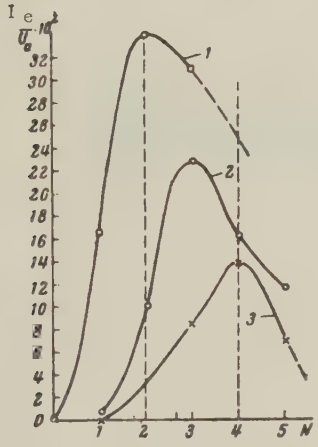


Figure 4. Dependence of the ratio I_e/U_a on the number of cycles of treating a cathoporetic cathode with oxygen at different coating thicknesses:
1, 2, 3 — l = 6, 20, 35 microns.

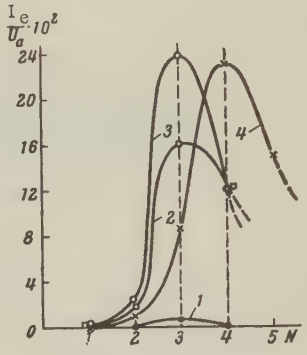


Figure 5. Dependence of I_e/U_a on the number of cycles of treating a sprayed cathode with oxygen at different coating thicknesses:
1, 2, 3, 4) l = 12, 20, 40, and 60 microns, respectively.

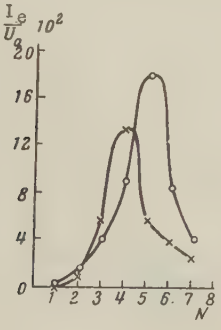


Figure 6. Dependence of I_e/U_a on the number of oxygen-treatment cycles for two cathodes with cores of platinum-coated nickel.

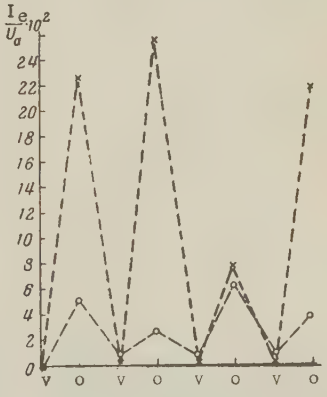


Figure 7. Variation of the ratio I_e/U_a with alternation of the cycles of heating the cathode in vacuum (v) and in oxygen (o).

completely loses its emitting ability after being subsequently heated in vacuum, and that the activation and deactivation process can be repeated an arbitrary number of times. Inasmuch as brief heating in vacuum at 850° cannot cause decomposition of the NiO, as is the case in most other oxides, the assumption that processing in oxygen leads to oxidation in the nickel base only, or to the oxidation of some impurities in the layers, can be regarded as inconsistent. Yet after repeated treatment of the cathode in oxygen, the thickness of the NiO layer produced between the core and the magnesium oxide layer should in all probability increase. The fact that after each annealling in vacuum the cathode can be restored to its previous degree of activity by heating in oxygen is evidence that neither the presence or absence of a nickel-oxide layer, nor the very thickness of the nickel oxide layer, influences the emitting properties of the cathode.

Recognizing that one can hardly expect a noticeable diffusion of the oxygen inside the magnesium oxide grains from the times and temperatures used in the heating in oxygen (850° C for three minutes), since it is known, for example, that the diffusion of oxygen in single crystals of magnesium oxide is observed only after heating to temperatures greater than 1500° for many hours), it is most likely that the treatment of the cathode in oxygen leads to the diffusion of the oxygen over the pores and to its adsorption on the surfaces of the magnesium oxide grains. Vacuum heating, to the contrary, leads to desorption of the oxygen and to its removal from the layer of magnesium oxide. Thus, the state of activity of the magnesium oxide cathode should be determined by the presence of a film of adsorbed oxygen on the surfaces of the magnesium oxide grains.

b. Temperature dependence of emission. The establishment of the character of the temperature dependence of the emission from a magnesium oxide cathode is of interest both from the point of view of elucidating the mechanism of emission, and from the point of view of determining the temperature stability of this cathode. An investigation of the temperature variation of the emission of an magnesium oxide cathode has shown that both reversible and irreversible changes in the emission current with temperature can be observed. Figure 8 shows the temperature variation of the emission current in the interval from 100 to 400°C. The curve discloses the presence of a rather sharply pronounced maximum, but the variations of the emission with temperature are fully reversible here.

Heating the cathode to higher temperatures leads to irreversible changes in emission, apparently resulting from the surface migration and desorption of oxygen (see Figure 9).

The decrease in emission with increasing temperature can explain the tendency to saturation, observed on the volt-ampere characteristics in the region of high current densities.

Direct experiment has shown that the temperature of the MgO layer (measured with a chromel-alumel thermocouple) greatly increases with increasing density of the current drawn, reaching values of current density on the order of several milliamperes per square centimeter, values corresponding to the decreasing branch of the emission current vs. temperature curve (see Figure 10). Thus, the limit on the increase in emission current density is imposed by

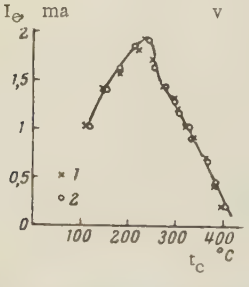


Figure 8. Temperature dependence of the emission in the region of reversible changes of the emission current (U_a constant):
1) increase in t_c; 2) decrease in t_c.

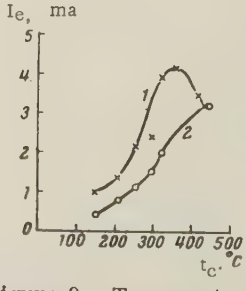


Figure 9. Temperature dependence of emission (case of irreversible change in emission current, U_a constant):
1) increase in t_c;
2) decrease in t_c.

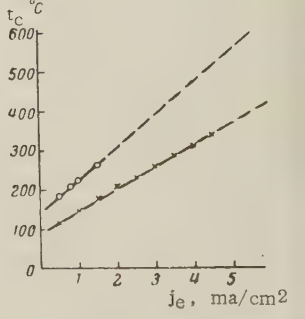


Figure 10. Dependence of temperature of magnesium oxide layer on the emission current density (for two different cathodes).

the heating of the magnesium oxide layer. Consequently the limiting current densities from an magnesium oxide cathode can be increased by using forced cooling.

c. Formation of emission. To explain the variation in the formation of the emission of the magnesium oxide cathode with time, we observed on an oscilloscope screen the variations of the emission current after turning on the anode voltage, with the starter electrode heater turned on beforehand. A typical oscillogram of the emission current is shown in Figure 11. The anode voltage was observed on the oscilloscope simultaneously with the emission current, and the step on the oscillogram corresponds to the instant when the anode voltage is turned on.

The oscillographs show that the duration of the formation process ranges between rather wide limits, depending on the applied anode voltage (which determines the steady-state stationary emission current) and, in particular, on the filament current of the starter electrode. Figure 12 shows the dependence of the formation time, determined from the oscillographs (that is, the time elapsed from the instant when the anode voltage is turned on to the establishment of the stationary current) on the density of the emission current established at the instant of the formation process (regulated by the anode voltage). The curve shown in Figure 12 shows that the formation time decreases from 0.5 to ~ 0.05 sec as the current density is increased from 1 to 7.5 ma/cm^2.

We deemed it interesting to ascertain which factor contributes to the greater degree to the formation of the emission, the electron emission of the starter electrode or the light

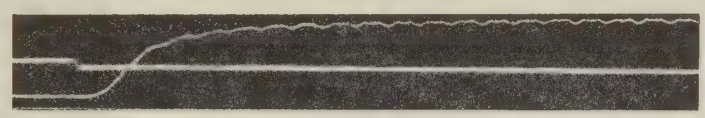


Figure 11. Oscillogram of formation of emission current upon application of the anode voltage.

radiated from it. For this purpose we measured the formation time with negative voltage (-60 v) applied to the starter electrode relative to the cathode (in the latter case, obviously, the electron current from the starter electrode to the cathode should be absolutely nil). The results of these measurements, illustrated by the curves of Figure 13 (plotted for two values of the steady-state emission current) indicate that with the emission current from the starter to the cathode cutoff, the formation of the emission is faster. It follows therefore, first, that for the formation of the emission it is sufficient only to illuminate the cathode, while bombardment with electrons under these conditions only makes the formation process more difficult (under the experimental conditions, when a negative voltage was applied to the starter electrode, the electrons emitted from this electrode bombarded the cathode until a trapping voltage was applied to the anode, and this apparently made it especially difficult to form the emission after turning on the anode voltage).

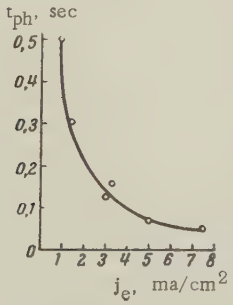


Figure 12. Dependence of the time of formation of emission on the density of the stationary emission current.

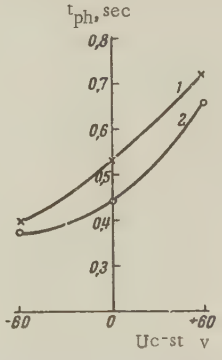


Figure 13. Dependence of the time of formation of emission on the potential of the cathode relative to the starter electrode for two values of the emission current. 1 — I_e = 3 ma, 2 — I = 10 ma

A reduction in the filament current of the starter electrode greatly increases the formation time. Thus, a reduction in the filament current of the starter from 1.5 to 0.9 amp has led to an increase in the emission formation time, other conditions being equal, from 0.4 to 1.7 sec.

Turning off the anode voltage causes practically instantaneous cessation of the emission current.

d. Effect of a magnetic field. Inasmuch as the magnesium oxide layers have a high porosity, an important factor in the emission may be the character of the motion of the electrons in the pores, which should change if a magnetic field is applied to the layer. To ascertain the degree of influence of a magnetic field on the emission, we measured the emission current at constant anode voltage as a function of the intensity of the magnetic field perpendicular to the surface of the cathode. The results of the measurements, illustrated by the curves of Figure 14, are evidence that the magnetic field actually does appreciably influence the emission, and an increase in the magnetic field intensity causes first an increase and then a reduction in the emission current.

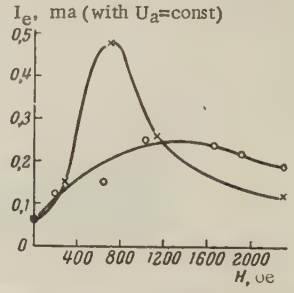


Figure 14. Dependence of emission current on the magnetic field intensity (for two cathodes).

e. Effect of thickness of the coating. As already noted, the thickness of the coatings in the cathode investigated vary considerably (from 6 to 35 microns in cathophoretic cathodes and from 12 to 16 microns in sprayed ones). For cathophoretic cathodes, as a rule, the quality of the cathodes improved with decreasing thickness of the coating; the optimum value of I_e/U_a increased (essentially owing to the decreased voltage U_a) (see Figure 4). The emission was more uniformly distributed over the surface of the cathode in this case, and vacuum annealing of thin coatings did not lead to the formation of cracks in the layers, whereas thick coatings always cracked (see Figure 1b).

f. Microscopic observations of the emitting surface. Self-maintaining emissions from a magnesium oxide cathode always is accompanied by luminescence of the emitting part of the surface. In the observation of a well activated (in oxygen) cathode with the unaided eye, the distribution of the luminescence of the surface of the cathode seems fully uniform. However, microscopic observations show that in the case of thick cathophoretic coatings, which become cracked after vacuum annealing, the luminescence is concentrated essentially in the regions of the cracks in the form of individual bright centers. For thin cathophoretic in the form of individual bright centers. For thin cathophoretic coatings the distribution of the luminescence over the surface of the cathode is found to be actually homogeneous. Heating of the cathode leads to the appearance of fluctuations in the cross section (with formation in some cases of fluctuating dark spots), which apparently is connected with migration and removal of oxygen from the layer.

It is interesting to note that when the current drawn is small, one observes sometimes that only part of the cathode surface glows, and the area of the glow increases with increasing current. This phenomenon recalls the well-known increase in the area of cathode glow with increasing current, observed in normal glow discharge.

g. On the excitation of self-maintained emission by direct application of voltage and on the role of the trapping electric field. If the role of the starter in the development of a self-maintained emission reduces to positively charging the surface of the magnesium oxide layer and to the production of an electric field in it, it is natural to expect that self-maintained emission can be produced by direct application of a voltage on the magnesium oxide layer. To verify this possibility we prepared a normally processed magnesium oxide cathode, to the surface of which we pressed a spring grid (wire diameter 50 microns, mesh 0.37 mm). A positive voltage was applied on the grid relative to the core of the cathode. We were unable, however, to excite emission in this manner, in spite of the fact that the voltage on the grid was raised to 800 v. At the same time, when the starter electrode was turned on, the emission was readily excited.

We can thus consider it established that in order for a self-maintaining emission to start developing it is essential to introduce into the layer a certain, albeit small, amount of free electrons. It is interesting to note that when a field is produced in the magnesium oxide layer

begins to break down at voltages around 750 v, whereas with the starter turned on the breakdowns were observed even at 300 v. This is obviously evidence of the unevenness in the distribution of the potential in the magnesium oxide layer under conditions when self-maintained emission exists.

To ascertain the role of the trapping field we prepared diodes whose mesh anode consisted of two parts, one a dense net and the other a loose one.* Experiment has shown that the emission was more readily formed in the part of the cathode located under the dense grid.

2. DISCUSSION OF RESULTS

The whole aggregate of the results of the research on self-maintaining emission from dielectric layers is quite obviously evidence that this effect is the result of a strong electric field present in the layer. What is debatable is which elementary processes, responsible for the development of the self-maintaining emission, are due to this field.

We consider it advisable to examine separately the phenomena in single crystals (and in solid bodies in general) and phenomena in porous systems.

The physics of semiconductors and dielectrics has established the following sequence of the main physical phenomena in continuous media as the electric field intensity E is increased. An increase in E produces, first, ionization of the impurities, following the Zener or Frenkel mechanism or, the ionization appears as a consequence of the increase in the number of fast electrons, due to the mechanism of impact ionization. Any further increase in E leads to an increase in overheating of the electron gas (that is, to an increase in its effective temperature), but the distribution function remains practially symmetrical. At a certain value of the effective electron temperature, a noticeable isotropic impact ionization of the lattice atoms begins, and this leads to trasitions of electrons from the filled band to the conduction band. At even stronger fields, in which the asymmetrical part of the distribution function f_1 becomes commensurate with its symmetrical part f_0, avalanchelike ionization develops. At large E (on the order of 10^7 v/cm), tunnel emission of the electrons from the metallic base into the dielectric layer is also possible.

The question arises whether development of a self-maintaining electron emission is possible in systems consisting of a continuous dielectric layer, deposited on a metallic base. Apparently such a possibility can not be refuted, but its mechanism cannot be connected with the development of an avalanche impact ionization over the entire thickness of the dielectric layer (which usually ranges from 1000 Å upward). The avalanchelike impact ionization in a solid can have a stationary character only in very thin layers, when the intensity of the avalanche is limited by the departure of the carriers from the region of the strong field (for example, in Reed avalanche diodes). At the same time one can visualize a situation wherein a highly inhomogeneous field distribution (see, for example, [9]) is established in the continuous dielectric (due to the ionization of the impurities) with the greater part of the potential drop occurring directly at the surface of the base, and this can lead to a development of an avalanche impact ionization only in the thin layer next to the contact, or also directly to tunnel emission from the metal into the dielectric without formation of an avalanche. In the remaining part of the dielectric layer, there can be established in this case a field which is insufficient for the development of the avalanche and can only maintain an effective electron temperature sufficiently high for emission and for avalanche-free ionization of the impurities, and consequently the process can acquire a stationary character. The possibility of a stationary self-maintaining emission from continuous dielectric layers is apparently confirmed by the earlier experiments with B_2O_3 films [6], and the possibility of tunnel emission from the metal into the dielectric layer with subsequent emission of the electrons into the vacuum without development of avalanche impact ionization in the dielectric layer follows from the experiments of M.I. Yelinson and A.G. Zhdan [10].

In porous systems, such as a magnesium oxide cathode, the mechanism of self-maintained emission apparently has a different character. Here it is possible for the electrons on become highly accelerated in the pores at fields which present no danger of breakdown (on the

---

* The data on the two batches of grids are as follows: wire diameter 0.03 mm, mesh 0.04 mm (dense) and 0.3 mm (loose); wire diameter 0.018 mm, mesh, 0.07 mm (dense) and 1.2 mm (loose).

order of 10^5 v/cm). Following references [1,2,11] we can assume that the process of development of the self-maintaining emission is connected in this case with formation of an electronic avalanche over the entire thickness of the layer, because primary electrons, formed deep in the layer knock from the grains of the MgO secondary electrons, which in turn are accelerated in the pores and participate in further development of the avalanche. The mechanism producing the primary electrons deep in the layer may be either the photoionization with quanta produced by recombination or possibly the tunnel effect from the base. The current is made stationary in this case by hole conductivity, arising in band-band impact ionization in the grains of the magnesium oxide.

From this point of view, the role of activation of the magnesium oxide layer with oxygen may reduce to the production of acceptor surface levels, the presence of which ensures a certain optimal hole conductivity of the MgO grains at room temperatures.

Luminescence of the layer in the visible part of the spectrum can apparently be attributed to recombination of oxygen centers. The temperature variation of the emission current (in the case of reversible changes) can be related both to the temperature variations of the hole conductivity, and to changes in the effective temperature of the electrons in MgO grains and the influence of the temperature on the departure of the electron from the grains into pores.

Evidence in favor of the avalanche impact ionization by electrons accelerated in the pores are, in addition to the relatively low mean values of the field in the layer (insufficient for the development of the avalanche impact ionization in the continuous medium; the relatively high mean energies of the emitted electrons ($\sim$13 ev), which exceed the threshold of impact ionization for magnesium oxide, and the possibility of obtaining self-maintaining emission from coatings of greater thickness.

At the same time these factors, as well as the independence of the emission on the presence and on the thickness of the intermediate NiO layer, raise doubts concerning the possibility of regarding tunnel emission from the base, with departure of the greater part of the electrons to the vacuum, as the main process in the mechanism of cold emission from a magnesium oxide cathode.

The clarification of this question will be the main object of future investigations.

The results of the present work show, as do the results obtained earlier by many other researchers, that the magnesium-oxide cathode is characterized by a relatively low limiting current density, very unfavorable form of distribution of the emitted electrons by energies (a broad spectrum ranging from 0 to 130 v) and a sufficiently long time of emission formation upon application of the anode voltage, excluding the possibility of working in the pulsed mode. These shortcomings limit to some extent the region of application of this cathode. At the same time, many advantages of cathodes with self-maintaining emission make them preferable for use in those electronic tubes, in which the foregoing shortcomings are not essential. In this connection, the organization of further research on cold magnesium oxide cathodes both for the purpose of obtaining more accurate ideas concerning its operating mechanism and for the purpose of improving its operating properties, is a timely problem.

REFERENCES

1. D. Dobischek, Electronics and Communs, 1959, 7, 5, 26.
2. A. M. Skellet, B. G. Firth, D. W. Mayer, Proc. I. R. E., 1959, 47, 10, 1704.
3. Y. Mizushima, Y. Igarashi, T. Imai, J. Phys. Soc. Japan, 1960, 15, 4, 729.
4. R. M. Aranovich, I. G. Ksendzatskiy, P. V. Timofeyev, Izvestiya AN SSSR, Division of Technical Science — Power and Automation 1960, 6, 143.
5. H. N. Daglish, Proc, I. E. E., 1961, 108B, 37, 103.
6. D. V. Zernov, Zh. Tekh. Fiz. 1937, 7, 1787.
7. D. V. Zernov, M. I. Yelinson, N. M. Levin, Izvestiya AN SSSR, Division of Technical Sciences, 1944, 3, 166.
8. R. M. Aranovich, ibid. 1944, 8, 346.
9. M. I. Yelinson, D. V. Zernov, Radiotekhnika i Elektronika 1957, 2, 1, 75.
10. M. I. Yelinson, A. G. Zhdan, ibid. 1959, 4, 1, 135.

Submitted to the Editors 23 May 1961

PENETRATION OF AN ELECTRIC FIELD INTO A SEMICONDUCTOR

G. F. Vasil'yev

Formulas are obtained describing the penetration of an external electric field into a semiconductor. The calculation is made under the assumption that the temperature of the electrons in a semiconductor is always equal to the lattice temperature.

1. The problem of the penetration of an electric field into a semiconductor was solved by Morgulis [1] and by Stratton [2]. Their results, however, are valid for two extreme cases, when the slope of the energy levels in the semiconductor ΔR is much less than kT[1] (k is Boltzmann's constant and T is the absolute temperature), and when $\Delta R + \zeta$ is much greater kT[2] (ζ is the Fermi energy). The most interesting intermediate region is not described by the formulas obtained in the foregoing papers. A formula valid for any energy interval ΔR is necessary in calculations of the field at which the degeneracy of the free carriers in the semiconductor takes place (degeneracy sets in when the bottom of the conduction band crosses the Fermi level). We derive below formulas for the penetration of the field in a semiconductor, valid for any energy interval ΔR.

2. Let us consider the one-dimensional case. The Poisson equation has the form

$$\begin{aligned} &\frac{d^2\varphi}{dx^2} = -\frac{4\pi\rho}{\varkappa}, \\ &\varphi(0) = \Delta R, \quad \left(\frac{d\varphi}{dx}\right)_{x=0} = \frac{F}{\varkappa}, \\ &\varphi(x) \to 0 \text{ where } x \to \infty, \\ &\frac{d\varphi}{dx} \to 0 \text{ where } x \to -\infty. \end{aligned} \tag{1}$$

Here F is the electric field; $\varkappa$ the volume charge; $\varkappa$ the dielectric constant. The volume charge consists of the charge $\pm en_{1,2}$ of the carriers (the upper sign pertains to the electrons or donors and the lower to the holes or acceptors) and the impurity-level charge $\pm en_{4,3}$ is the concentration of the ionized atoms of the donor impurities n_4 or acceptor impurities n_3. The charges on the different impurity levels and the numbers of the carriers are in definite equilibrium relationships with each other. The number of charges having the same charge as the carriers is directly proportional to the number of the latter, and their presence in accounted for by multiplying $n_{1,2}$ by a factor P, greater than unity, depending on the concentration of the binding levels and on their position relative to the bottom of the conduction band. On the other hand, the number of charges of opposite sign is determined from the equilibrium condition. If the number of donors or acceptors per unit volume is N_4 or N_3, we can obtain by the law of effective masses

$$n_{4,3} = \frac{N_{4,3}}{1 + C_1 n_{1,2}}. \tag{2}$$

The volume charge is

$$\rho = \mp e^2 \left(P n_{1,2} - \frac{N_{4,3}}{1 + C_1 n_{1,2}} \right), \tag{3}$$

where

$$C_1 = \frac{N_{4,3} - P n_{1,2}^{\infty}}{P n_{1,2}^{\infty 2}}; \tag{4}$$

$n_{1,2}^{\infty}$ is the concentration of the carriers away from the surface of the semiconductor. The value of $n_{1,2}$ can be calculated with the aid of the Fermi integral

$$n_{1,2} = 4\pi \frac{(2m_{1,2}kT)^{3/2}}{h^3} \int_0^{\infty} \frac{\sqrt{x}\, dx}{1 + e^{x - \beta_{1,2}}} = C_2 F_{1/2}(\beta_{1,2}), \tag{5}$$

where

$$C_2 = 4\pi \frac{(2m_{1,2}kT)^{3/2}}{h^3};$$

$$F_{1/2}(\beta_{1,2}) = \int_0^\infty \frac{\sqrt{x}\,dx}{1+e^{x-\beta_{1,2}}};$$

$\beta = \zeta/kT$; $x = \varepsilon/kT$; ε is the electron energy; the subscripts "1" and "2' pertain respectively to the electron and hole semiconductors. After integrating by parts and breaking up the integration limits, we can rewrite the Fermi integral

$$2F_{1/2}(\beta_{1,2}) = \int_0^{\beta_{1,2}} x^{-1/2}(\beta - x)\,dx + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \int_0^{\beta_{1,2}} e^{-ny}(\beta - y)^{-1/2} + \\ + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \int_0^{\infty} e^{-ny}(\beta + y)^{-1/2}dy. \qquad (6)$$

Here $y = \beta - x$.

$\beta \gg 1$, we have

$$(\beta - y)^{-1/2} \simeq \beta^{-1/2},$$

In this case we obtain

$$(\beta + y)^{-1/2} \simeq \beta^{-1/2}.$$

$$F_{1/2}(\beta_{1,2}) \simeq \frac{2}{3}\beta_{1,2}^{3/2} + \frac{\pi^2}{12}\beta_{1,2}^{-1/2}. \qquad (7)$$

When $\beta \ll 1$ the exponential factor predominates, and therefore

$$F_{1/2}(\beta_{1,2}) \simeq \frac{\pi^2}{10}\ln(1 + e^{\beta_{1,2}}). \qquad (8)$$

A function more convenient for integration can be obtained in the form

$$\frac{\pi^2}{11}\frac{e^{\beta_{1,2}}}{1+\frac{e^{\beta_{1,2}}}{4}} \simeq \frac{\pi^2}{10}\ln(1 + e^{\beta_{1,2}}). \qquad (9)$$

The table lists the numerical values of the resultant functions, and also the value of the Fermi integral, taken from the tables of Stoner and McDougall [3]. We measure the Fermi energy in an electronic semiconductor from the bottom of the conduction band upward, while in a hole-type semiconductor we measure it from the top of the filled band downward.

Using the function (9) and also Equations (1) and (3) in the region $\beta \leqslant 2$, we obtain

$$F^2 = 8\pi e^2 \varkappa kT \left\{ PC_2 \ln \frac{4+\exp\left(\frac{\zeta+\Delta R}{kT}\right)}{4+\exp\left(\frac{\zeta}{kT}\right)} - N_{4,3}\frac{\Delta R}{kT} + \right. \\ \left. + \frac{4\frac{\pi^2}{11}C_2C_1N_{4,3}}{1+4\frac{\pi^2}{11}C_2C_1} \ln \frac{4+\left(1+4\frac{\pi^2}{11}C_2C_1\right)\exp\left(\frac{\Delta R+\zeta}{kT}\right)}{4+\left(1+4\frac{\pi^2}{11}C_2C_1\right)\exp\left(\frac{\zeta}{kT}\right)} \right\}. \qquad (10)$$

β	$\frac{\pi^2}{10}\ln(1+e^{\beta})$	$\frac{\pi^2}{11}\frac{e^{\beta}}{1+e^{\beta/4}}$	$\frac{2}{3}\beta^{3/2}\frac{\pi^2}{12}\beta^{-1/2}$	$F_{1/2}(\beta)$
—5	0.0067	0.0060	—	0.0060
—4	0.017	0.0161	—	0.0161
—3	0.049	0.0439	—	0.0434
—2	0.128	0.1165	—	0.1146
—1	0.306	0.301	—	0.2905
0	0.683	0.7168	—	0.6781
1	1.290	1.4515	—	1.3964
2	—	2.325	2.46	2.5025
3	—	2.99	3.93	3.977

Formula (10) gives the result of Morgulis and Stratton [1,2] if $\Delta R \ll kt$ and $\exp(\zeta/kT) \ll 1$.

$$\frac{\Delta R}{kT} \simeq 1 - \exp\left(-\frac{\Delta R}{kT}\right),$$

$$\ln \frac{4+\exp\left(\frac{\zeta+\Delta R}{kT}\right)}{4+\exp\left(\frac{\zeta}{kT}\right)} \simeq \frac{1}{4}\exp\left(\frac{\zeta}{kT}\right)\left[\exp\left(\frac{\Delta R}{kT}-1\right)\right];$$

$$\ln \frac{4+\left(1+4\frac{\pi^2}{11}C_2C_1\right)\exp\left(\frac{\Delta R+\zeta}{kT}\right)}{4+\left(1+4\frac{\pi^2}{11}C_2C_1\right)\exp\left(\frac{\zeta}{kT}\right)} \simeq 0.$$

In the numerical coefficients, putting $m_1 = m_2 = m$ (m is the mass of the free electron), we can rewrite (10) in the form

$$F^2 = 3.08\cdot 10^{-10}\varkappa T\left\{2.23\cdot 10^{16}T^{3/2}P\ln\frac{4+\exp\left(\frac{\zeta+\Delta R}{kT}\right)}{4+\exp\left(\frac{\zeta}{kT}\right)} - N_{4.3}\frac{\Delta R}{kT} + \right.$$

$$\left. + \frac{2.23\cdot 10^{16}T^{3/2}C_1N_{4.3}}{1+2.23\cdot 10^{16}T^{3/2}C_1}\ln\frac{4+(1+2.23\cdot 10^{16}T^{3/2}C_1)\exp\left(\frac{\Delta R+\zeta}{kT}\right)}{4+(1+2.23\cdot 10^{16}T^{3/2}C_1)\exp\left(\frac{\zeta}{kT}\right)}\right\}\left(\frac{\text{ev}}{\text{cm}}\right)^2.$$

Degeneracy sets in when $\Delta R = -\zeta$. For n-germanium at T=300°K, we get $N_4 = 10^{14}\,\text{cm}^{-3}$; if P = 1, degeneracy sets in at $F = 5\cdot 10^6$ ev/cm. Formula (10) is valid up to $\Delta R = -\zeta + 2kT$.

When $\beta > 2$, (1), (3), and (7) yield

$$F^2 = 8\pi e^2\varkappa\, kTPC_2\left[\frac{4}{15}\left(\frac{\Delta R+\zeta}{kT}\right)^{5/2} + \frac{\pi^2}{6}\left(\frac{\Delta R+\zeta}{kT}\right)^{1/2} - 2.08\right]. \qquad (11)$$

In the derivation of (11) we have neglected the charge on the impurity levels. Using numerical values we can represent (11) in the form (if $m_1 = m_2 = m$)

$$F^2 = [2.7\cdot 10^{16}\varkappa\,(\Delta R+\zeta)^{5/2} + 1.22\cdot 10^9\varkappa T^2\,(\Delta R+\zeta)^{1/2} -$$

$$- 1.43\cdot 10^7\varkappa T^{5/2}]\left(\frac{\text{ev}}{\text{cm}}\right)^2.$$

If we neglect in (11) the last two two terms in the square brackets, we obtain Stratton's results [2].

REFERENCES

1. N.D. Morgulis, Zh. Eksp. Teoret. Fiz, 1948, 16, 11, 959.

Submitted to the Editors 7 September 1960

DETERMINATION OF THE LIFETIME OF NONEQUILIBRIUM CURRENT CARRIERS IN SEMICONDUCTORS FROM THE SPECTRAL DISTRIBUTION OF THE NOISE

L. Ya. Pervova

From the spectral distribution of the noise we determine the lifetimes of the nonequilibrium carriers. We also determine the corresponding capture cross section from the known concentration of the recombination centers. The cross section of the capture of a hole by a singly charged gold atom in germanium was found to be $(2.3 - 3.1)\cdot 10^{-14}$ cm^2.

It was indicated in [1] that an investigation of the spectral distribution of the noise in semiconductors makes it possible to determine the lifetime of the nonequilibrium majority carriers.

Actually, we can regard it as established that in the region of low and medium frequencies the noise spectrum can be represented by an expression having the following form (see, for example, [2]):

$$F(\omega) = \frac{A}{\omega^m} + \frac{B}{1+(\omega\tau)^2} \quad (m \simeq 1), \tag{1}$$

where the first term is the noise component called in the literature the "1/f noise" (we shall not discuss its mechanism here), while the second term is the generation-recombination (shot) noise, due to the fluctuations in the carrier concentration in the semiconductor.

At low frequencies the $1/\omega$ noise predominates, but with increasing frequency it rapidly decreases until at a certain frequency $\omega_1 = 2\pi f_1$ this component become negligibly small compared with the shot noise. If this frequency is such that we still have $\omega\tau \ll 1$, then the spectral density of the noise becomes independent of the frequency. Finally, at frequencies $\omega \to 1/\tau$ the noise curve begins to drop off and $F(\omega)$ becomes inversely proportional to the square of the frequency. At the point where $F(\omega) = \frac{1}{1+(\omega\tau)^2} = \frac{1}{2}$, we have

$$\omega\tau = 1. \tag{2}$$

This leads to a direct determination of the lifetime τ.

In the case where for some reason or other it is difficult to reach the decreasing branch of the shot noise, we can use for the determination of the lifetime τ the well-known expression for the shot noise

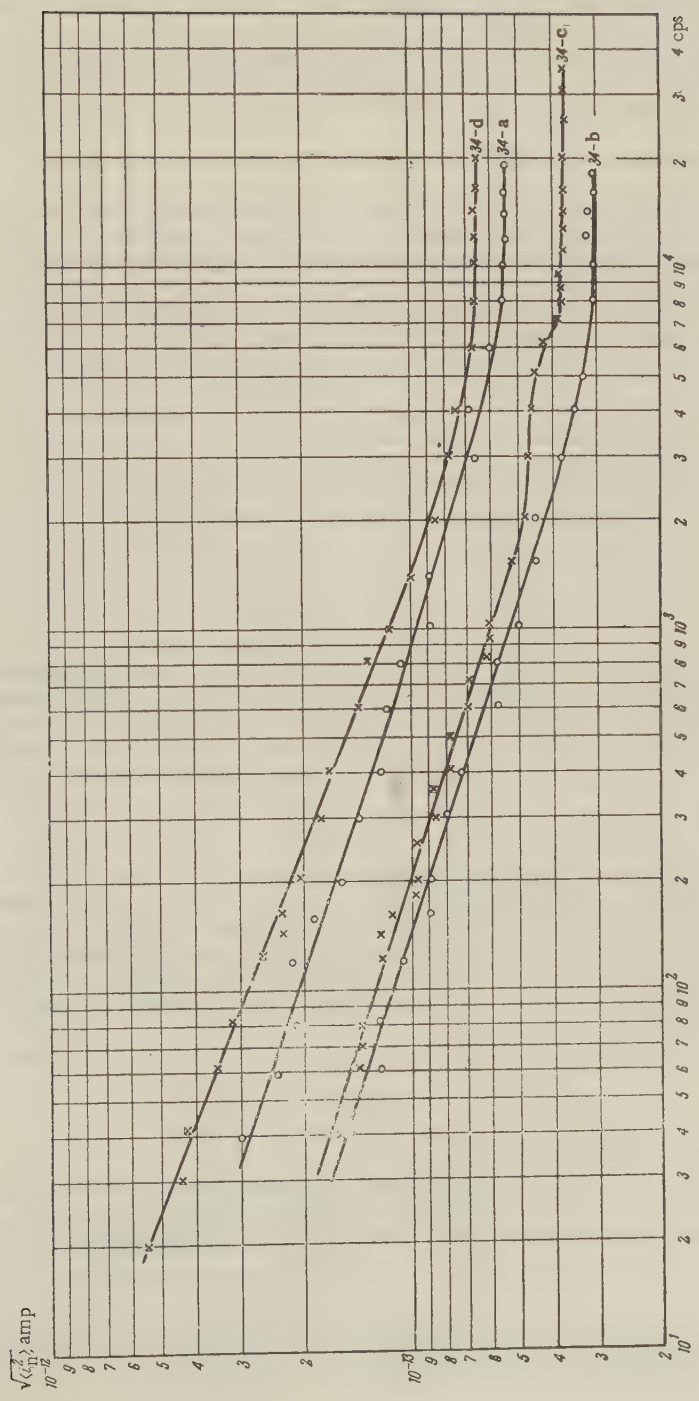
$$\langle \overline{i_n^2} \rangle = 4eI \frac{\tau}{\tau_L} \frac{1}{1+(\omega\tau)^2} \Delta f, \tag{3}$$

where $\tau_L = L/\mu E$ is the time that the carrier needs to cover a path equal to the length L of the specimen; μ is the mobility of the carriers.

This expression is valid if the energy gap ΔE between the level and the band with which the carriers are exchanged is much greater than kT.

Knowing the carrier mobility μ and measuring $\langle i_n^2 \rangle$ in the region of the plateau, where the spectral density of the noise does not change with frequency, we can determine the lifetime τ from the measurements or, vice versa, we can determine τ from condition (2) and then determine the mobility of the carriers by measuring $\langle i_n^2 \rangle$.

The method described for determining the lifetime τ has, it appears to us, at least two obvious advantages: first, it makes it possible to measure τ under conditions of rather small deviations from equilibrium (low "injection" levels), and second, it permits a relatively simple measurement of small lifetimes τ even in cases when they amount to only tenths or even hundredths of a microsecond.


$\sqrt{\langle i_n^2 \rangle}$ amp
10^{-12}
10^{-13}
cps
34-d
34-a
34-c
34-b

We note also that this method is founded on rather general thermodynamic considerations, which are the basis of the calculation of the shot noise, and is therefore apparently applicable in all cases when it is possible to separate the shot noise experimentally.

By way of example we give below the results of measuring the lifetimes of holes in p-type germanium, alloyed with gold and compensated with antimony to various degrees.

The measurements were carried out on an audiofrequency analyzer with a bandwidth $\Delta f_{ef} = 4$ cps. The noise of the specimens was amplified with an amplifier having a gain about $2 \cdot 10^5$. The entire setup was calibrated against the noise from a vacuum antimony-cesium photocell. Special measurements have shown that this noise is white over the entire range of frequencies employed.

The measured specimens were placed in a cryostat where they were kept at a temperature of 80°K.

The spectral characteristics of the noise of four specimens, referred to one volt applied to the specimen, are shown in the figure. As can be seen from the figure, the shot noise manifests itself clearly starting with frequencies $f > 8 \cdot 10^3$ cps. The higher limit of 20,000 cps of the analyzer did not permit us to observe with these specimens the decreasing branch of the shot noise (specimen 34-c was measured up to 35,000 cps with an AS-3 analyzer).

Calculations based on formula (3) taking into account the carrier mobility, specimen length L = 6 cm, and the current flowing through the sample, I, leads to the lifetime values listed in the table.

Sample no.	N_{Au}, cm^{-3}	N_{Sb}, cm^{-3}	μ, $\frac{cm^2}{v\text{-}sec}$	I, per volt μa	τ, sec	σ, cm^2
34- d	$1.81 \cdot 10^{15}$	$5.6 \cdot 10^{13}$	$1.3 \cdot 10^4$	2.66	$5.75 \cdot 10^{-8}$	$3.1 \cdot 10^{-14}$
34- b	$2.5 \cdot 10^{15}$	$6.9 \cdot 10^{13}$	$1.05 \cdot 10^4$	1.1	$5.50 \cdot 10^{-8}$	$2.6 \cdot 10^{-14}$
34- a	$1.65 \cdot 10^{15}$	9.10^{13}	$2.2 \cdot 10^4$	1.96	$4.2 \cdot 10^{-8}$	$2.6 \cdot 10^{-14}$
34- c	$4.25 \cdot 10^{15}$	$2.02 \cdot 10^{14}$	$1.68 \cdot 10^4$	2.22	$2.1 \cdot 10^{-8}$	$2.34 \cdot 10^{-14}$

The table lists the concentration of gold N_{Au} and of the compensating impurity antimony N_{Sb}.*

Assuming that the given values of τ are connected with the recombination of holes with singly charged centers of the gold, the number of which is N_{Au}^- is obviously equal to N_{Sb}, we obtain the effective cross section for the recombination of holes on these centers.

$$\sigma_{Au^-} = \frac{1}{N_{Au}^- v \tau},$$

where $v = 10^7$ cm/sec is the thermal velocity of the holes.

Numerical values of the recombination cross sections are listed in the last column of the Table. These values are close to those found by Rupprecht [3], namely $\sigma_{Au} = 5.7 \cdot 10^{-14} cm^2$.

In conclusion, it should be pointed out that by investigating the noise spectrum of samples illuminated with monochromatic light of different wavelengths, it is possible to determine the characteristic times and the corresponding effective cross sections for different impurity centers in semiconductors.

The author is grateful to V.N. Makarov who participated in the measurements.

REFERENCES

1. L.Ya. Pervova, Radiotekhnika i Elektronika, 1959, 4, 2, 330.
2. A. van der Ziel, Fluctuations in Radio and in Physics (Russian Translation) GEI, 1958, page 140.
3. G. Rupprecht, Proc. Internat. Conf. Semic. Phys., Prague, 1960, 282.

Submitted to the Editors 24 January 1961

---

*The author is grateful to V.G. Alekseyeva for preparing the specimens and to V.I. Sidorov for measuring the concentration of the impurities and the mobility of the carriers in these samples.

PRECISION TRANSISTOR AMPLIFIERS FOR ELECTRONIC ANALOGS WITH HIGH INPUT (OR OUTPUT) RESISTANCE

Yu. L. Kurkin, N. S. Kurkina

Methods are proposed for the design of transistor circuits which provide high input r_{in} and (or) output r_{out} resistances in current or voltage amplifiers and in voltage to current converters. It is shown that the high output resistance of such circuits, $r_{out} \gg r_c$, can be obtained only by using positive feedback. Specific circuits are described, and some of their practical characteristics are given. It is shown how to connect the bias circuits so as not to shunt the high resistances r_{in} and r_{out}.

In electronic analog computer technology use is made of precision voltage amplifiers with high input resistances r_{in}, or of current amplifiers with very high output resistances r_{out} [1, 2]. In the design of transistor amplifiers, a very important factor is the reduction in the sensitivity of the parameters of the amplifier to changes in the parameters of the active elements. Stabilization is ensured by using feedback. In particular, current amplifiers, which should have a low input and high output resistance, are stabilized by parallel-series feedback.

In existing circuits of current feedback amplifiers (Figure 1), the output resistance is limited to the value of the collector resistance r_{c3} of the output transistor, which shunts the output for any value of negative feedback. Actually, according to the limit as $\alpha \to 1$ in the expression for r_{out} (Table 1, circuit of Figure 1), we obtain $r_{out} \to r_{c3}$.

A high output resistance $r_{out} \gg r_c$ is obtained by introducing slight positive feedback. The circuit shown in Figure 2 is distinguished by a very high output resistance, $r_{out} \gg r_{c3}$, exceeding by many times the resistance of the collector of the output transistor. A characteristic feature of this circuit is the simultaneous use of over-all feedback and local positive feedback. Such a multiple-loop feedback increases the stability of the circuit parameters and decreases the sensitivity of these parameters to changes in the properties of the active elements [3]. When the negative feedback loop is open [resistance (1-h) R], the positive-feedback loop, which includes T_1 and T_2, is also opened. All that remains is the feedback from the base to the emitter of T_3 through the transistor T_2. The amount of this feedback does exceed $\alpha_2(1-\alpha_3) \ll 1$, and therefore the problem of stability does not arise.

The input resistance r_{in}, the output resistance r_{out}, and the current gain of the circuit of Figure 2 are listed in Table 1.

It can be shown that there exists an optimum value of positive feedback. Indeed, when

$$h = h_{opt} = \frac{(1-\alpha_2)\left[1+(1-\alpha_1)\frac{R}{R_e}\right]}{\alpha_1\alpha_2\frac{R}{R_e}} \simeq \frac{1}{20} \qquad (1)$$

the denominator of r_{out} vanishes and $r_{out} \to \infty$, and $A_1 \to 1 + R/R_e$. It is easy to see that h_{opt} is a function of the temperature and of the operating mode.

Assume that $h = h_{opt} - \Delta h$. Then

$$r_{out} = r_{c3}\frac{\alpha_3 + (1-\alpha_3)\,h\,\frac{\Delta h}{h}\,\frac{R}{R_e}\,\alpha_2}{h\,\frac{\Delta h}{h}\,\frac{R}{R_e}\,\alpha_2}. \qquad (2)$$

Table 1

Parameters of Current Amplifiers With High Output Resistance

1) Circuit	A_i	r_{out}	r_{in}
Fig. 1	$\dfrac{\alpha_1\alpha_3\left(1+\dfrac{R}{R_e}\right)}{\alpha_1+\left(1+\dfrac{R}{R_e}\right)(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}$	$r_{C3}\dfrac{\alpha_1+\left(1+\dfrac{R}{R_e}\right)(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}{\alpha_1+\left(1+\dfrac{R}{R_e}\right)(1-\alpha_1)(1-\alpha_2)} \leqslant r_{C3}$	$\dfrac{{}_1L_{11}(1-\alpha_2)(1-\alpha_3)\left(1+\dfrac{R}{R_e}\right)}{\alpha_1+\left(1+\dfrac{R}{R_e}\right)(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)} \leqslant 0{,}05\,\text{ohm}$
Fig. 2	$\dfrac{1+\dfrac{R}{R_e}}{1+\dfrac{1-\alpha_3}{\alpha_1\alpha_3}\left\{(1-\alpha_2)\left[1+\dfrac{R}{R_e}(1-\alpha_1)\right]-L\dfrac{R}{R_e}\alpha_1\alpha_2\right\}}$	$\dfrac{\alpha_1\alpha_2 r_{C3}\left\{1+\dfrac{1-\alpha_3}{\alpha_1\alpha_3}\left\{(1-\alpha_2)\left[1+\dfrac{R}{R_e}(1-\alpha_1)\right]-L\dfrac{R}{R_e}\alpha_1\alpha_2\right\}\right\}}{(1-\alpha_2)\left[1+\dfrac{R}{R_e}(1-\alpha_1)\right]-L\dfrac{R}{R_e}\alpha_1\alpha_2}$	$\dfrac{{}_1L_{11}(1-\alpha_2)(1+\alpha_3)\left(1+\dfrac{R}{R_e}\right)}{\alpha_1\alpha_3+(1-\alpha_2)(1-\alpha_3)-L\dfrac{R}{R_e}\alpha_1\alpha_2(1-\alpha_3)}$
Fig. 2 $h=h_{opt}$	$1+\dfrac{R}{R_e}$	$\simeq\dfrac{r_{C3}}{(1-\alpha_2)}\dfrac{1}{\dfrac{\Delta L}{L}}$	$\dfrac{{}_1L_{11}(1-\alpha_2)(1-\alpha_3)\left(1+\dfrac{R}{R_e}\right)}{\alpha_1\alpha_3+\dfrac{R}{R_e}(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}$

Substituting in (2) $h = h_{opt}$, we obtain

$$\frac{r_{out}}{r_{cs}} = \frac{\alpha_1\alpha_3 + (1-\alpha_2)(1-\alpha_3)\left[1+\frac{R}{R_e}(1-\alpha_1)\right]\frac{\Delta h}{h}}{(1-\alpha_2)\left[1+\frac{R}{R_e}(1-\alpha_1)\right]\frac{\Delta h}{h}}, \qquad (3)$$

$$r_{out} = \frac{r_{c3}}{1-\alpha_2}\,\frac{1}{\frac{\Delta h}{h}}.$$

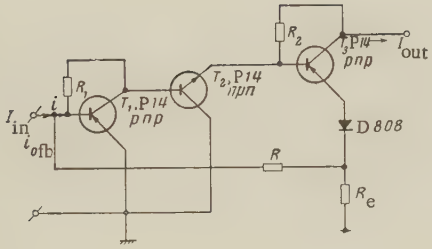


Figure 1. Diagram of a current amplifier with feedback; the output resistance is limited by the shunting action of r_c.

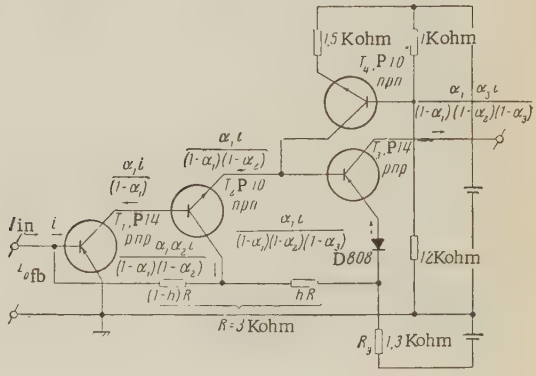


Figure 2. Diagram of a current amplifier with mixed (positive and negative) feedback, providing high output resistance $r_{out} \gg r_c$.

In practice

$$\frac{\Delta h}{h} = 10 - 20\% = -\left\{\frac{\Delta\alpha_1}{\alpha_1}\,\frac{1+\frac{R}{R_e}}{1+\frac{R}{R_e}(1-\alpha_1)} + \frac{\Delta\alpha_2}{\alpha_2}\,\frac{\alpha_1}{1-\alpha_2}\right\},$$

that is, r_{out} = 50 to 300 meg.

When $h = h_{opt}$ we have

$$Z_{in} = Z_{in}^{*} = \frac{{}_1h_{11}(1-\alpha_2)(1-\alpha_3)\left(1+\frac{R}{R_e}\right)}{\alpha_1\alpha_3 + (1-\alpha_1)(1-\alpha_2)(1-\alpha_3)\frac{R}{R_e}}.$$

The stability of the current gain is

$$\frac{\Delta A_i}{A_i} = \frac{\Delta\alpha_1}{\alpha_1}\,\frac{(1-\alpha_2)(1-\alpha_3)\left(1+\frac{R}{R_e}\right)}{\alpha_1\alpha_3} +$$

$$+\frac{\Delta\alpha_2}{\alpha_2}\,\frac{1-\alpha_3}{\alpha_1\alpha_3}\left[1+\frac{R}{R_e}(1-\alpha_1)\right]\left[1-(1-\alpha_2)\frac{\Delta h}{h}\right] +$$

$$+\frac{\Delta\alpha_3}{\alpha_3}\,\frac{\frac{\Delta h}{h}(1-\alpha_2)\left[1+\frac{R}{R_e}(1-\alpha_1)\right]}{\alpha_1\alpha_3}.$$

It is easy to see that introduction of optimal positive feedback increases the stability of the circuit. When $h = h_{opt}$ the gain of the circuit is independent of α_3.

The expression for r_{out} (Table 1) can be obtained by substituting in the expression for A_1 instead of α_3 its dynamic value $\alpha_3/(1+Z_L/r_{c3})$. Then A_1 can be represented in the form

$A_{ic3}/(1 + Z_L/r_{out})$

The return difference can be obtained for the circuit shown in Figure 2 by opening the negative feedback circuit (resistance R[1 - h]), and connecting the free end of R(1 - h) to ground. Taking into account the dependence of α_1 and α_2 on R and R_e, which determine the input resistance of the second and third stages and the output resistance of the first stage $h_{11}/(1 - \alpha_1)$, we can show that for a specified R/R_e there exists an optimum value of the resistances R and R_e ensuring a maximum return difference. In practice $R_{opt} = 0.5 - 1$ k when $R/R_e = 1$.

An examination of the form of the expressions for the parameters of the current amplifier (Table 1) shows that an investigation of the circuits aimed at obtaining a high output resistance should be carried out in the following sequence.

a) The current gain $A_1 = A_1(Z_L)$ is derived. The numerator of A_1 contains the factor α_3, the current transfer of the output transistor; α_3 depends on the load, namely $\alpha_3 = \alpha_{30}/(1 + Z_L/r_{c3})$. This circumstance indeed limits in ordinary circuits the value of r_{out} to the value of r_{c3}.

b) We derive similar terms of the denominator, containing α_3.

c) We impose the condition that all the terms of denominator which do not contain α_3 be set equal to zero, something possible only if the denominator contains terms with minus signs, introduced by the positive feedback circuit including the output stages. With this, the values of α_3 of the numerator and denominator cancel out. Consequently, r_{c3} no longer limits the output resistance of the circuit. It follows therefore that to obtain high output resistance of the output transistor ($r_{out} \gg r_c$), one must use only circuits with positive feedback including the output stages. Certain experimental characteristics of the circuit of Figure 2 are shown in Figures 3a, b.

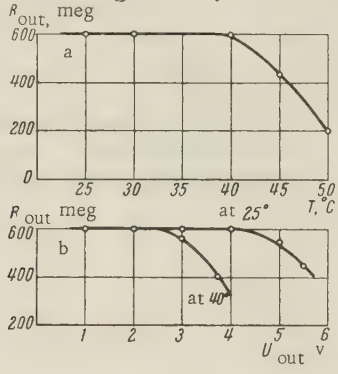


Figure 3. Dependence of R_{out} on the temperature (a) and on the output voltage (b).

A high input resistance in a transistor voltage amplifier can be obtained with only a single feedback loop (parallel-series). Attention should be paid to prevent the high input resistance from being shunted by the bias circuits and by the collector resistances r_{c1}, r_{c2} etc. The bias circuits (resistors R_1, R_2, etc., see Figure 4) are connected in parallel with the base — collector junctions of the corresponding transistors. To eliminate leakage through the parallel connections $r_{c1}R_1/(r_{c1} + R_1)$, $r_{c2}R_2/(r_{c2} + R_2)$ etc., the corresponding collectors should be connected to part of the output voltage, equal to or nearly equal to the input voltage, either directly or through small input resistances of the succeeding stages.

Figure 4 shows the diagram of voltage amplifier with a high input resistance $r_{in} \gg r_c$ (on the order of 500 — 2000 meg):

$$r_{in} = \frac{\alpha_1\alpha_2\alpha_3}{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)} \times \cdot$$

$$\times \frac{1}{\left[\frac{1}{R_e} + \frac{1}{R_3(1-\alpha_3)}\right] + \frac{1}{r_{c3}(1-\alpha_3)}} + \left(1 + \frac{R}{R_c}\right)\left[\frac{1}{Z_L} + \frac{1}{r_{e4}(1-\alpha_4)}\right] \cdot$$

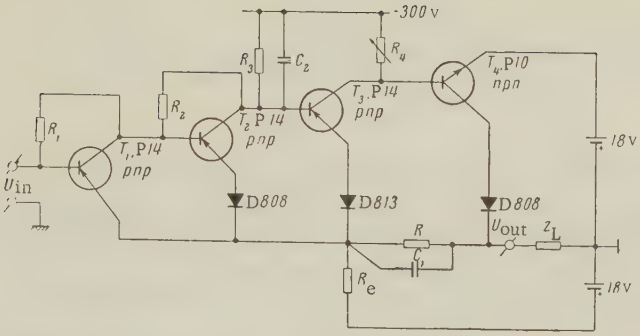


Figure 4. Diagram of an amplifier with high input resistance.

The total input resistance is the parallel combination of the following three resistances:

$$\frac{\alpha_1\alpha_2\alpha_3 R_e}{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)(1-\alpha_4)}; \quad \frac{\alpha_1\alpha_2\alpha_3 (R_3 \parallel r_{c3})}{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)(1-\alpha_4)};$$

$$\frac{\alpha_1\alpha_2\alpha_3 r_{c4}}{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)\left(1+\frac{R}{R_e}\right)}.$$

The stability of the voltage gain under changes in the operating condition, temperature, or upon replacement of the transistors is ensured by the series-parallel negative feedback, which increases the input resistance and decreases the output resistance. Inasmuch as the

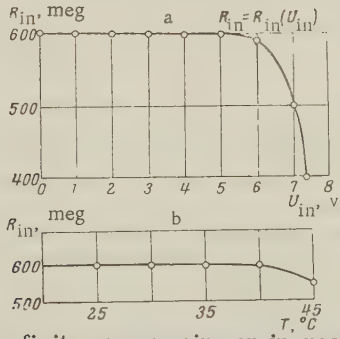


Figure 5. Practical characteristics of the amplifier circuit shown in Figure 4; the dependence of the input resistance on the temperature (a) and on the signal level (b).

transistor has a finite power gain, an increase in the input resistance is possible only at the expense of the voltage gain. To be sure, the voltage gain is thereby stabilized. The voltage gain is

$$A_V = \frac{\frac{1}{R_e} + \frac{1}{R} + \frac{1}{r_{c3}(1-\alpha_3)} + \frac{1-\alpha_4}{\alpha_4 R}\left[1 + \frac{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}{\alpha_1\alpha_2\alpha_3}\right]}{\frac{1}{R} + \frac{1-\alpha_4}{\alpha_4}\left[\frac{1}{R} + \frac{1}{Z_L} + \frac{1}{r_{c4}(1-\alpha_4)}\right]\left[1 + \frac{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}{\alpha_1\alpha_2\alpha_3}\right]},$$

$$A_V = \frac{\left(1+\frac{\alpha_4 R}{R_e}\right) + \frac{\alpha_4 R}{r_{c3}(1-\alpha_3)}}{1 + \frac{(1-\alpha_4)R}{Z_L} + \frac{R}{r_{c4}}} = 1 + \frac{\alpha_4 R}{R_e}.$$

Resistances R_1 and R_2 do not shunt the input, since their free ends are connected through the low-resistance input circuits of transistors T_1 and T_2 to the voltage U_1, which is practically equal to the input voltage. In the absence of feedback from the last stage, the current through R_e is $\alpha_1\alpha_2\alpha_3/(1-\alpha_1) \times (1-\alpha_2)(1-\alpha_3)$ times larger than the current at the input. Consequently, R_e yields an input resistance $\sim\alpha_1\alpha_2\alpha_3 R_3/(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)$. The parallel-series feedback increases r_{in} by another $1/(1-\alpha_3)$ times. It is easy to see (Figure 4) that $R_e(1-\alpha_3)$ and $R_3(1-\alpha_3)$ and $r_{c3}(1-\alpha_3)$ are connected in parallel to R_e. The output

resistance of the circuit is

$$r_{out} = \frac{(R + R_e)\left[R_L + \frac{RR_e}{R + R_e}\frac{\alpha_1\alpha_2\alpha_3}{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}\right]}{R_L + \frac{\alpha_1\alpha_2\alpha_3 R_e}{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)(1-\alpha_4)}}.$$

Figures 5a and b show some experimental characteristics of the voltage amplifier.

Let us consider now a transistor amplifier with high input and output resistances, many times greater than the collector resistance r_c of the transistors employed. Amplifiers of this type are called voltage — current converters or ideal pentodes [4].

The input and output resistances are increased by introducing series feedback. In existing circuits, this increase is limited by the value of r_c.

The circuit which we propose, shown in Figure 6, makes it possible to overcome this

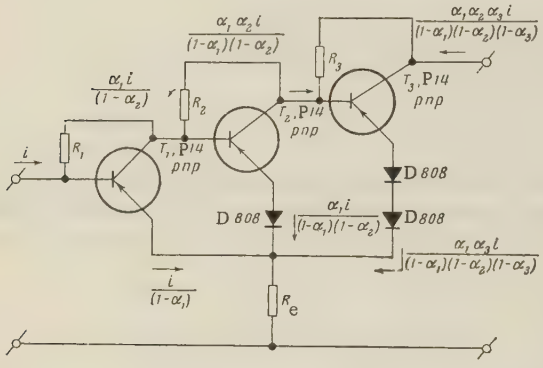


Figure 6. Diagram of a voltage—current converter with high input and output resistances.

limitation, owing to the combined utilization of a common negative feedback and a local positive feedback.

Table 2

Parameters of High-Resistance Voltage — Current Converter

Transconductance
$S = \frac{I_{out}}{U_{in}} = \frac{\alpha_1\alpha_2\alpha_3\left[\frac{1}{R_e} + \frac{1}{1-\alpha_3}\left(\frac{1}{r_{c3}} + \frac{1}{R_3}\right)\right]}{\alpha_1\alpha_3 + (1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}$
Input resistance
$r_{in} = \frac{\alpha_1\alpha_2\alpha_3 + (1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}{\frac{(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)}{R_e} + (1-\alpha_1)(1-\alpha_2)\left(\frac{1}{r_{c3}} + \frac{1}{R_3}\right)}$
Output resistance (r_{out}) with input short circuited
$r_{out} = \frac{(r_{c3} \parallel R_3)[\alpha_1\alpha_2\alpha_3 + (1-\alpha_1)(1-\alpha_2)(1-\alpha_3)]}{(1-\alpha_1)(1-\alpha_2)}\left[\frac{\frac{1}{R_e} + \frac{1}{r_{c3}} + \frac{1}{R_3}}{\frac{1}{R_e} + \frac{1}{1-\alpha_3}\left(\frac{1}{r_{c3}} + \frac{1}{R_3}\right)}\right]$

Table 2 lists some parameters of the circuit shown in Figure 6.

In these formulas r_{c3} must be taken to include the parallel resistance R_3, which feeds the output transistor.

All the formulas have been derived under the assumption that the external resistances are $R \gg h_{11}$.

Taking into account the input resistances of the transistor proper, $h_{11}/(1-\alpha)$, leads to a slight increase in the obtained high input or output resistances.

The expressions derived agreed will with experiment within the limits determined by the changes in the transistor parameters with temperature and with operating conditions.

REFERENCES

1. Yu. L. Kurkin, N. S. Kurkina, Precision transistor integrator, Avtomatika i Telemekhanika, 1961, 22, 7, 907.
2. Edward M. Davis, Comparisons between multiple loop and single loop transistor feedback amplifiers, I. R. E. Wescon Convention Record, 1959, p. 78-86.
3. M. Kawakami, Some fundamental considerations on active four-terminal linear networks. I. R. E. Trans., 1958, CT-5, 2, 115.
4. Yu. L. Kurkin, A. A. Sokolov, Negative impedance converter using transistors, Elektrichestvo, 1959, 9, 66.

Submitted to the Editors 3 November 1960,
after revision 23 July 1961

BRIEF COMMUNICATIONS

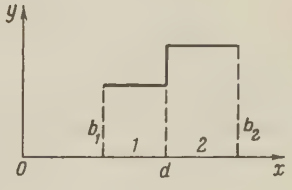
CONNECTION BETWEEN THE METHOD OF UNDETERMINED COEFFICIENTS AND THE INTEGRAL-EQUATION METHOD

A. Ya. Yashkin

In many works [1, 2, 3] on the design of waveguides, the method of undetermined coefficients [4] is applied, and it is noted that this method is preferable to the method of integral equations [5], since it is connected as it were with the approximation of an unknown function. In fact the method of undetermined coefficients, like the other method, is based on the approximation of an unknown function, but one chooses as the approximated function always a function through which one specifies the initial fields, whereas in the solution of the integral equation one can use for the approximating functions arbitrary functions, which comprise a complete system on the given interval.

We shall show below that the method of undetermined coefficients is a particular case of the integral-equation method, namely a case in which we choose as the approximating functions for the solution of the equation such functions, in terms of which the initial fields are specified. As a particular case, the method of undetermined coefficients is naturally less useful than the general method and consequently one cannot speak of any comparison of this method to the integral-equation method.

The correctness of the foregoing premise can be demonstrated by making direct calculations by both methods. We shall confine ourselves to an examination of a rectangular Cartesian system. It can be seen, however, that an analogous derivation can be made [6] in any other system of coordinates in which the variables separate, since the orthogonality of the eigenfunctions of the wave equation in the Cartesian system, which we use here, remains in force for the wave equations in other systems that admit of separation of variables.



Let a waveguide carrying an H mode have a steplike form in cross section. We break up the steplike section into a series of rectangular regions, and choose two neighboring regions, which we denote by the numbers 1 and 2 (see the figure). We write for each region the general solution of the wave equation for the axial component $H_Z = H(x,y)$ in the form of an infinite sum of the particular solutions corresponding to one and the same eigenvalue k:

$$H_1 = C_{10}\cos k(x-\alpha_1) + \sum_{m=1}^{\infty} C_{1m}\cos r_{1m} y \operatorname{ch} r'_{1m}(x-\alpha_1), \tag{1}$$

where

$$H_2 = C_{20}\cos k(x-\alpha_2) + \sum_{n=1}^{\infty} C_{2n}\cos r_{2n} y \operatorname{ch} r'_{2n}(x-\alpha_2), \tag{2}$$

$$r_{1m} = \frac{m\pi}{b_1}; \quad r_{2n} = \frac{n\pi}{b_2}; \quad m = 1, 2, \ldots; \quad n = 1, 2, \ldots;$$

$$r'_{1m} = \sqrt{r_{1m}^2 - k^2}; \quad r'_{2n} = \sqrt{r_{2n}^2 - k^2};$$

α_1 and α_2 are parameters determined from the conditions of the continuity of the solutions on the interface between the individual rectuangular regions (these are not significant in our calculations).

The continuity conditions on the line x=d

$$H_1 = H_2, \quad \frac{\partial H_1}{\partial x} = \frac{\partial H_2}{\partial x} \tag{3}$$

yield two equations

$$C_{10}\cos k(d-\alpha_1) + \sum_{m=1}^{\infty} C_{1m}\cos r_{1m} y \operatorname{ch} r'_{1m}(d-\alpha_1) = \tag{4}$$

$$= C_{20}\cos k(d-\alpha_2) + \sum_{n=1}^{\infty} C_{2n}\cos r_{2n} y \operatorname{ch} r'_{2n}(d-\alpha_2),$$

$$-C_{10}k\sin k(d-\alpha_1) + \sum_{m=1}^{\infty} C_{1m} r'_{1m} y \operatorname{sh} r'_{1m}(d-\alpha_1) =$$

$$= -C_{20}k\sin k(d-\alpha_2) + \sum_{n=1}^{\infty} C_{2n} r'_{2n}\cos r_{2n} y \operatorname{sh} r'_{2n}(d-\alpha_2) = \varphi(y) \sim E_y. \tag{5}$$

The problem now consists of obtaining from (4) and (5) a relation between the number k and the dimensions of the two regions. We obtain these relations by two methods.

First method (method of integral equation). We consider expression (5) as an expansion of an unknown function $\varphi(y)$ in a Fourier series in $\cos r_{1m} y$ in the interval $(0, b_1)$ and in $\cos r_{2n}(y)$ in the interval $(0, b_2)$, and we express the coefficients C_{1m} and C_{2n} in terms of $\varphi(y)$. Substituting them into the expression (4) we obtain (assuming $b_2 > b_1$) the homogeneous integral equation

$$\int_0^{b_1} \varphi(\xi)\,[P(y,\xi) + P_0]\,d\xi = 0, \tag{6}$$

in which the symmetrical kernel is given by the expressions

$$P(y,\xi) = \sum_{m=1}^{\infty} \frac{2\operatorname{cth} r'_{1m}(d-\alpha_1)}{r'_{1m} b_1}\cos r_{1m} y \cos r_{1m}\xi -$$

$$- \sum_{n=1}^{\infty} \frac{2\operatorname{cth} r'_{2n}(d-\alpha_2)}{r'_{2n} b_2}\cos r_{2n} y \cos r_{2n}\xi, \tag{7}$$

$$P_0 = \frac{\operatorname{ctg} k(d-\alpha_2)}{kb_2} - \frac{\operatorname{ctg} k(d-\alpha_1)}{kb_1}.$$

Applying to solution (6) the Bubnov-Galerkin method (or the Ritz variational method), we obtain the sought relation, which connects the unknown number k with the dimensions of the two regions

$$\begin{vmatrix} M_{00}+P_0N_{00}, & M_{01}+P_0N_{01}\ldots & M_{0i}+P_0N_{0i}\ldots \\ M_{10}+P_0N_{10}, & M_{11}+P_0N_{11}\ldots & M_{1i}+P_0N_{1i}\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots & & \\ M_{j0}+P_0N_{j0}, & M_{j1}+P_0N_{j1}\ldots & M_{ji}+P_0N_{ji}\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots & & \end{vmatrix} = 0 \tag{8}$$

where

$$M_{ji} = \int_0^{b_1}\int_0^{b_1} \varphi_j(\xi)\, P(y, \xi)\, \varphi_i(y)\, d\xi\, dy;$$

$$i = 0, 1, 2, \ldots; \; j = 0, 1, 2, \ldots$$

$$N_{ji} = \int_0^{b_1}\int_0^{b_1} \varphi_j(\xi)\, \varphi_i(y)\, d\xi\, dy; \tag{9}$$

φ_j and φ_j are approximating functions, with the aid of which the unknown function $\varphi(\xi)$ is represented in series form

$$\varphi(\xi) = \sum_{i=0}^{\infty} \beta_i \varphi_i(\xi). \tag{10}$$

Second method (method of undetermined coefficients). We introduce in (4) and (5) the notation

$$C_{10}\cos k(d-\alpha_1) = B_0, \qquad C_{20}\cos k(d-\alpha_2) = D_0,$$

$$C_{1m}\operatorname{ch} r'_{1m}(d-\alpha_1) = B_m, \qquad C_{2n}\operatorname{ch} r'_{2n}(d-\alpha_2) = D_n, \tag{11}$$

$$-C_{10}k\sin k(d-\alpha_1) = \beta_0, \qquad -C_{20}k\sin k(d-\alpha_2) = A_0,$$

$$C_{1m}r'_{1m}\operatorname{sh} r'_{1m}(d-\alpha_1) = \beta_m, \qquad C_{2n}r'_{2n}\operatorname{sh} r'_{2n}(d-\alpha_2) = A_n, \tag{12}$$

after which these equations become

$$B_0 + \sum_{m=1}^{\infty} B_m \cos r_{1m}y = D_0 + \sum_{n=1}^{\infty} D_n \cos r_{2n}y = f(y), \tag{13}$$

$$\beta_0 + \sum_{m=1}^{\infty} \beta_m \cos r_{1m}y = A_0 + \sum_{n=1}^{\infty} A_n \cos r_{2n}y = \varphi(y). \tag{14}$$

According to the method of undetermined coefficients, we must express the coefficients A_n, B_m, and D_n in terms of β_m.

Regarding A_0 and A_n as Fourier coefficients of the expansion of $\varphi(y)$ in cos $r_{2n}(y)$ in the interval $(0, b_2)$, we obtain (recognizing that $\varphi(y)$ is equal to zero on the segment b_2-b_1)

$$A_0 = \frac{1}{b_2}\int_0^{b_1} \varphi(y)\, dy, \qquad A_n = \frac{2}{b_2}\int_0^{b_1} \varphi(y)\cos r_{2n}y\, dy. \tag{15}$$

Substituting the series (14) in the resultant expression in lieu of $\varphi(y)$ (this precisely is the gist of approximating the unknown function by means of functions in terms of which a field is specified in the region with the smaller dimension b), we express the coefficients A_n in terms of β_m.

The coefficients $B_j (j = 0, 1, \ldots)$ are determined from (13) in terms of the function f(y):

$$B_0 = \frac{1}{b_1}\int_0^{b_1} f(y)\, dy, \qquad B_j = \frac{2}{b_1}\int_0^{b_1} f(y)\cos r_{1j}y\, dy. \tag{16}$$

Instead of f(y) we substitute the sum from (13) with coefficients D_n; we obtain

$$B_0 = D_0 + \sum_{n=1}^{\infty} \frac{\sin r_{2n}b_1}{r_{2n}b_1} D_n, \tag{17}$$

$$B_j = \frac{2}{b_1}\sum_{n=1}^{\infty} \frac{r_{2n}\sin r_{2n}b_1}{r_{2n}^2 - r_{1j}^2} D_n. \tag{18}$$

But from relations (11) and (12), and also from (15) and (16), we can find that

$$B_0 = -\frac{\operatorname{ctg} k(d-\alpha_1)}{k}\beta_0; \quad B_j = \frac{\operatorname{cth} r'_{1j}(d-\alpha_1)}{r'_{1j}}\beta_j, \tag{19}$$

$$D_0 = -\frac{\operatorname{ctg} k(d-\alpha_2)}{k}A_0 - \frac{b_1}{b_2}\frac{\operatorname{ctg} k(d-\alpha_2)}{k}\beta_0, \tag{20}$$

$$D_n = \frac{\operatorname{cth} r'_{2n}(d-\alpha_2)}{r'_{2n}}A_n = \frac{2}{b_2}\frac{\operatorname{cth} r'_{2n}(d-\alpha_2)}{r'_{2n}}\frac{\sin r_{2n}b_1}{r_{2n}}\beta_0 + \\ + \frac{2}{b_2}\frac{\operatorname{cth} r'_{2n}(d-\alpha_2)}{r'_{2n}}\sum_{m=1}^{\infty}\frac{r_{2n}\sin r_{2n}b_1}{r_{2n}^2 - r_{1m}'^2}\beta_m. \tag{21}$$

Substituting the resultant values of the coefficients B, D in the expressions (17) and (18), we obtain (with variable i) a system of homogeneous equations linear in β

$$\sum_{j=0}^{\infty} L_{ji}\beta_i = 0; \quad i = 0, 1, 2, \ldots, \tag{22}$$

where L_{ji} are the coefficients of β_1 and are determined from the relations

$$L_{ji} = \frac{b_1}{2}\frac{\operatorname{cth} r'_{1j}(d-\alpha_1)}{r'_{1j}} - \frac{2}{b_2}\sum_{n=1}^{\infty}\frac{\operatorname{cth} r'_{2n}(d-\alpha_2)(r_{2n}\sin r_{2n}b_1)^2}{r'_{2n}(r_{2n}^2 - r_{1j}^2)(r_{2n}^2 - r_{1j}^2)}; \quad j \neq 0; \quad i \neq 0,$$

$$L_{j0} = -\frac{2}{b_2}\sum_{n=1}^{\infty}\frac{\operatorname{cth} r'_{2n}(d-\alpha_2)}{r'_{2n}}\frac{\sin r_{2n}b_1}{(r_{2n}^2 - r_{1j}^2)}; \quad j \neq 0; \quad i = 0,$$

$$L_{0i} = -\frac{2}{b_2}\sum_{n=1}^{\infty}\frac{\operatorname{cth} r'_{2n}(d-\alpha_2)}{r'_{2n}}\frac{(\sin r_{2n}b_1)^2}{r_{2n}^2 - r_{1i}^2}; \quad j = 0, \quad i \neq 0,$$

$$L_{00} = \frac{b_1}{b_2}\frac{\operatorname{ctg} k(d-\alpha_2)}{k} - \frac{\operatorname{ctg} k(d-\alpha_1)}{k} - 2\frac{b_1}{b_2}\sum_{n=1}^{\infty}\frac{\operatorname{cth} r'_{2n}(d-\alpha_2)}{r'_{2n}}\left(\frac{\sin r_{2n}b_1}{r_{2n}b_1}\right)^2; \\ j = i = 0.$$

Writing down the conditions for obtaining a nontrivial solution of (22), we obtain the sought equation relating the unknown quantity k with the dimensions of the two rectangular regions:

$$\begin{vmatrix} L_{00}, & L_{01}, & \ldots & L_{0i}\ldots \\ L_{10}, & L_{11}, & \ldots & L_{1i}\ldots \\ \ldots\ldots\ldots\ldots\ldots \\ L_{j0}, & L_{j1}, & \ldots & L_{ji}\ldots \\ \ldots\ldots\ldots\ldots\ldots \end{vmatrix} = 0. \tag{23}$$

We have thus obtained by different methods two equations, (8) and (23), which are solutions of one and the same problem. Let us show now that Equation (23) is obtained from (8) if we choose approximating functions that describe the field in the rectangular region with the lower height b, that is, if the unknown function is represented by the series

$$\varphi(\xi) = \sum_{i=0}^{\infty}\beta_i \cos r_{1i}\xi,$$

where $r_{1j} = i\frac{\pi}{b_1}$; $i = 0, 1, \ldots$. With this representation of $\varphi(\xi)$, all the N_{jr} except N_{00} vanish. Evaluation of the integrals of (9) leads to the following result:

$$M_{ji} + P_0 N_{ji} = \frac{b_1}{2}\frac{\operatorname{cth} r'_{1j}(d-\alpha_1)}{r'_{1j}} - \frac{2}{b_2}\sum_{n=1}^{\infty}\frac{\operatorname{cth} r'_{2n}(d-\alpha_2)(r_{2n}\sin r_{2n}b_1)^2}{r'_{2n}(r_{2n}^2 - r_{1j}^2)(r_{2n}^2 - r_{1i}^2)} = L_{ji},$$

$$M_{j0} + P_0 N_{j0} = L_{j0},$$

$$M_{0i} + P_0 N_{0i} = L_{0i},$$

$$M_{00} + P_0 N_{00} = L_{00},$$

which proves the assumption made above.

REFERENCES

1. D.V. Golovin, Propagation of Lower Modes in Certain Homogeneous and Inhomogeneous Waveguides, Dissertation, Moscow Extension Institutes for Communications, 1950.
2. N.F. Funtova, On Approximate Methods of Calculating the Critical Frequencies of π-waveguides. Scientific Notes of the V.I. Lenin Moscow State Polytechnic Institude, 1954, 88, 67.
3. V.M. Sedykh, Investigation of Waveguides With Cross-Shaped and H-Shaped Cross Sections, Dissertation, Khar'kov State University, 1959.
4. L.V. Kantorovich, V.I. Krylov, Approximate Methods of Higher Analysis, GTI, 1952.
5. L.N. Deryugin, Calculation of Critical Frequency of π and H-Waveguides, Radiotekhnika, 1948, 3, 6.
6. A.Ya. Yashkin, Concerning One Method of Calculating Straight and Bent Waveguides of Complicated Cross Section in Systems Admitting of Separation of Variables. Radiotekhnika i Elektronika, 1961, 6, 5, 754.

Submitted to the Editors 24 February 1961

ADJUSTABLE DIRECTIONAL COUPLER WITH CONSTANT PHASE SHIFT

M.Ya. Mandel'shtam

INTRODUCTION

In microwave techniques one uses adjustable directional couplers with slotted coupler [1,2]. A distinguishing feature of the principal diagram of such a coupler (Figure 1) is the fact that the electrical length of one of the two waveguides, connecting the slotted coupler, is adjustable. The adjustment makes it possible to change the transfer attenuation of the directional coupler over a wide range. However, simultaneously with changing the transfer attenuation, a change takes place in the phase shift of the wave passing through the directional coupler.

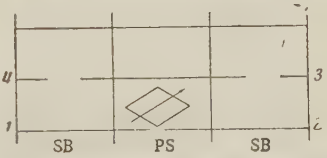


Figure 1. Principal diagram of adjustable directional coupler with one phase shifter: SB slotted coupler, PS phase shifter

In many cases it is useful to be able to regulate the transfer attenuation without dependence on the "parasitic" phase shift. Thus, for example, a ring resonator is best tuned to the optimum operating mode [3] if the change in the transfer attenuation does not influence the electrical length of the ring. A device permitting regulation of the amplitudes of the microwave field without affecting the phase would be of interest in phase measurements.

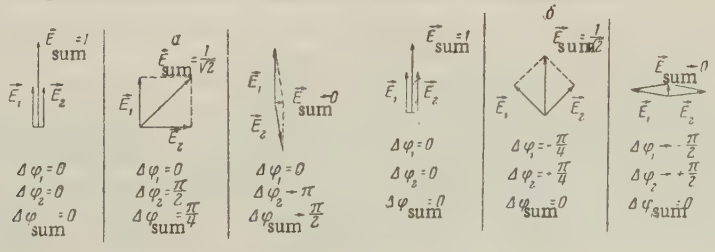


Figure 2. Vector diagrams of adjustable directional couplers: a) with one phase shifter; b) with two phase shifters.

The additional phase shift can be eliminated by simultaneously changing the lengths of both waveguides connecting the slotted couplers, in such a way that these changes are the same absolute magnitude but opposite in direction:

$$\Delta\varphi_1 = -\Delta\varphi_2, \tag{1}$$

where $\Delta\varphi_1$ and $\Delta\varphi_2$ are the increments of the electrical lengths of the first and second waveguides joining the slotted couplers.

The foregoing can be illustrated with the vector diagrams shown in Figure 2. On these diagrams $\vec{E}_1$ corresponds to a wave passing through the first waveguide, joining the slotted couplers; $\vec{E}_2$ corresponds to the wave transmitted through the second waveguide; $\vec{E}_{sum} = \vec{E}_1 + \vec{E}_2$ is the total field intensity, in a definite cross section of the output arm of the directional coupler (for example, in arm 2 or 3, Figure 1). A zero phase shift corresponds to the orientation of the vector $\vec{E}_{sum}$ when the vectors $\vec{E}_1$ and $\vec{E}_2$ are in phase; $\Delta\varphi_1$, $\Delta\varphi_2$, $\Delta\varphi_{sum}$ are respectively the phase shifts of vectors $\vec{E}_1$, $\vec{E}_2$, and $\vec{E}_{sum}$ relative to the zero phase.

These diagrams, which are shown in Figure 2, enable us to draw the following conclusions.

1. If the adjustment of the transfer attenuation is effected by changing the electrical length of one of two waveguides connecting the slotted couplers (usual directional coupler, Figure 1), then as the transfer attenuation changes from 0 to ∞ the phase of the outgoing wave changes by $\pi/2$.

2. If the regulation of the transfer attenuation is by means of simultaneous changes of the electrical length of both waveguides, equal in magnitude but opposite in direction, then the phase of the outgoing wave does not change in such an adjustment.

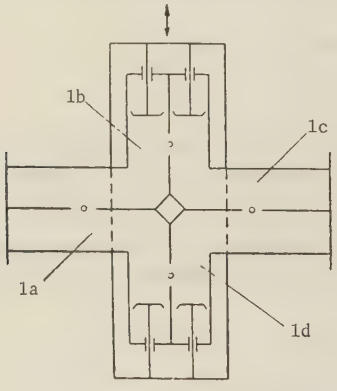


Figure 3. Principal diagram of an adjustable directional coupler with two rigidly coupled plunger-type phase shifters

1. ADJUSTABLE DIRECTIONAL COUPLER WITH PLUNGER-TYPE PHASE SHIFTERS

As is well known [1], the plunger-type phase shifter results in a transmitted-wave phase shift which is directly proportional to the displacement of the plunger block:

$$\Delta\varphi = 4\pi\frac{\Delta x}{\lambda_g}, \tag{2}$$

where $\Delta\varphi$ is the phase shift, Δx the change in the coordinate determining the position of the plunger block, and λg is the wavelength in the waveguide.

This circumstance enables us to design very simply an adjustable directional coupler with constant phase shift (Figure 3). All the plungers in this device are rigidly coupled to each other. When the upper pair of plungers is brought closer to the slotted coupler 1b, the lower pair of the plungers moves away from the slotted coupler 1d. With this, the electrical length of the upper waveguide, joining the slotted couplers 1a and 1c, decreases by a certain amount, and the electrical length of the lower waveguide increases by exactly the same amount; thus the condition that the phase shift of the total outgoing wave be constant is maintained.

2. ADJUSTABLE DIRECTIONAL COUPLER WITH FERRITE PHASE SHIFTERS

An analogous device can obviously be designed with the aid of ferrite phase shifters.

In this design, the phase shifters used are rectangular waveguides, partially filled with ferrite, and placed in the transverse magnetic field of an electromagnet.

Figure 4 shows the experimentally obtained dependence of the phase shift on the current in the magnetizing coil. An examination of Figure 4 shows that over a sufficiently wide range (between points a and b) the phase shift is in practice directly proportional to the magnetizing current. Analogous data are given in the literature [4,5].

As can be seen from Figure 2b, to change the transfer attenuation from 0 to ∞, with a constant phase shift, it is necessary to change uniformly the electrical length of one of the waveguides connecting the slotted couplers, from 0 to $+\pi/2$, and to change the other from 0 to $-\pi/2$. This requirement can be satisfied, by moving simultaneously and uniformly along the characteristic of one of the phase shifter from the point b to the point a.

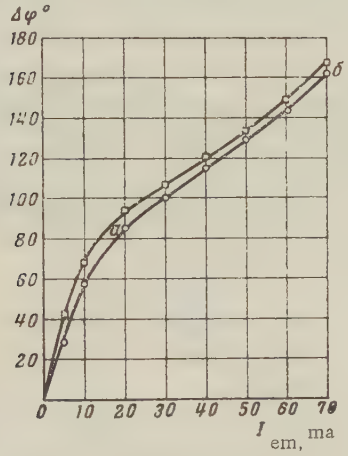


Figure 4. Dependence of the phase shift on the current in the magnetizing coil for two ferrite phase shifters

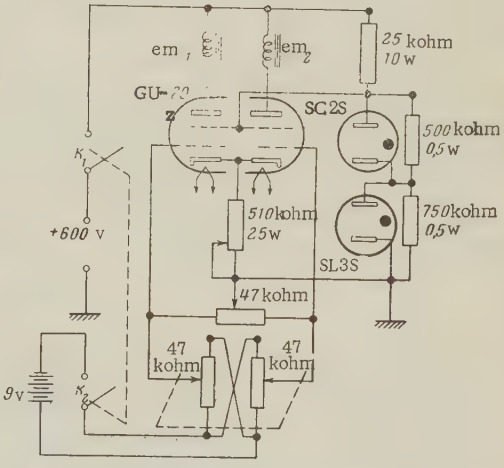


Figure 5. Principal electrical diagram of the supply for the electromagnets of the ferrite phase shifters

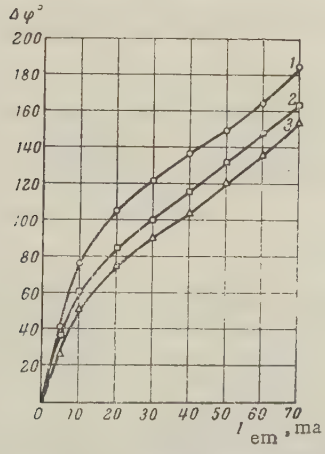


Figure 6. Dependence of the phase shift on the current in the magnetizing coil at different wavelengths:

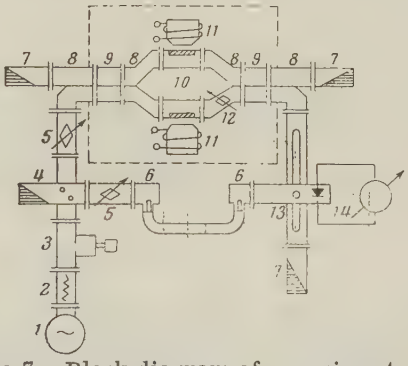


Figure 7. Block diagram of experimental setup: 1) generator; 2) ferrite gate; 3) wave meter; 4) directional coupler; 5) attenuator; 6) coaxial-waveguide junction; 7) absorber; 8) single to double waveguide junction; 9) slotted coupler; 10) ferrite phase shifter; 13) measuring line — phase meter; 14) microammeter (the dashed line surrounds the elements contained in the adjustable directional phase shifter with constant phase shift)

Figure 5 shows a simple electrical circuit which makes it possible to effect the aforementioned regulation of the phase shifts. When the voltage on the control grid of the left tetrode is increased by a certain amount, the voltage on the grid of the right tetrode decreases by the same amount. Accordingly the plate currents, which flow through the magnetizing coils (EM_1 and EM_2) of the electromagnets change accordingly. The difference in the characteristics in the right and left tetrodes can be used to compensate for the difference in the characteristics of the phase shifters. The use of tetrodes reduces the required stability of the plate-voltage source.

Measurements carried out in a range $\Delta\lambda/\lambda = \pm 3\%$, have shown that the slope of the portion of the curve of phase shift vs. magnetizing current between points a and b is practically independent of the wavelength (Figure 6), so that the adjustable directional coupler can be used in the indicated waveband.

An experimental check made with a laboratory model (Figure 7) has shown that in the wavelength range $\Delta\lambda/\lambda = \pm 3\%$ the phase shift remains practically constant, accurate to $\pm 2.5°$, as the transfer attenuation changes from 1 to 15 or 20 db. At the same time, when the transfer attenuation is regulated with the aid of one of the phase shifters (with the second phase shifter disconnected), a phase shift of $\sim 90°$ is observed in the output wave (see diagram, Figure 2a).

An interesting feature of the device considered is the invariance of the differential phase shift (that is, the difference in phase shifts for the two propagation directions) to regulation of the transfer attenuation. For the chosen dimensions of the ferrite plates, the differential phase shift was close to 180° and as the transfer attenuation was varied in the forward direction it changed by not more than $\pm 2°$ in the range from 1 to 15 or 20 db.

Thus, the apparatus considered has the properties not only of an adjustable directional coupler, but also of a gyrator [6]. This quality of the directional coupler can be used, in particular, to suppress backward waves in a traveling-wave resonator [7].

REFERENCES

1. L. Milosevic, R. Vautey, Le Vide, 1956, 65, 410.
2. W. L. Teeter, K. R. Bushore, I. R. E. Trans., 1957, MTT-5, 227.
3. V. E. Golant, M. Ya. Mandel'shtam, Radiotekhnika i Elektronika, 1959, 4, 4, 660.
4. R. F. Soohoo, I. R. E. Convention Record, 1956, pt. 5, 84-98.
5. M. L. Kales, H. N. Chait, N. G. Sakiotis, J. Appl. Phys., 1953, 24, 816.
6. A. D. Fox, S. E. Miller, M. T. Weiss, Properties of Ferrites and Their Use in Microwaves (Russian Translation), Soviet Radio Press, 1956.
7. K. Tomiyasu, I. R. E. Trans., 1957, MTT-5, 267.

Submitted to the Editors 3 May 1961

"PERTURBATION" OF A RESONATOR IN THE H_{01n} MODE BY A CYLINDRICAL ROD

E. B. Zal'tsman, V. E. Poyarkova

In many microwave measurements it is necessary to use narrowband (high-Q) systems, which can be frequency tuned with a high degree of resolution. Usually such a system is an electromagnetic resonator operating at the H_{01n} mode (this being the mode with highest Q), the frequency of which is adjusted with the aid of a piston. It is obvious that the resolving power (the piston displacement per unit frequency) is small in this case.

Relatively long ago, H_{01n} resonators have been described tunable with the aid of a cylindrical rod inserted through the sidewall; these resonators have a high resolving power (so-called semicoaxial wave meters [1]). There were no calculations, however, for such a resonator.

To calculate the change in the resonant frequency of the resonator upon insertion of a cylindrical rod through the end wall (see Figure 1), use is made of the well-known perturbation formula for the electromagnetic resonator {see, for example, [2], formula (P.26)}.

Substituting into this formula the well-known expressions for electric and magnetic fields of the H_{01} wave [2] and integrating in a cylindrical coordinate system we obtain the following formula for the frequency detuning of the resonator:

$$\frac{f-f_0}{f_0}=\frac{\gamma}{4\pi n J_0^2(\xi_{01})}\left(\frac{r}{R}\right)^2\left\{x(\beta_2-\beta_1)+\sin x\left[\left(\frac{2}{\gamma}-1\right)\beta_1-\beta_2\right]\right\}, \qquad (1)$$

where f_0 is the resonant frequency of the unperturbed resonator; f is the same with the rod inserted;

$$\gamma=\left(\frac{\xi_{01}}{2\pi}\right)^2\left(\frac{\lambda}{R}\right)^2;$$

$$x=4\pi\frac{h}{\lambda_B};\quad \beta_1=J_0^2(y)+J_1^2(y)-\frac{2}{y}J_1(y)J_0(y);$$

$$\beta_2=J_0^2(y)+J_1^2(y);\quad y=\xi_{01}\frac{r}{R};$$

J_0 and J_1 are Bessel functions of the first kind of zero and first order; ξ_{01} the first root of the function $J_1(y)$, $\xi_{01}=3.8317$; λ the TEM wavelength corresponding to the frequency f_0. The quantities r, R, h, and L are defined in Figure 1. It is tacitly assumed that $L=n\lambda_W/2$, where λ_W is the length of the periodicity along the resonator axis; n is the number of half-waves along the axis of the resonator.

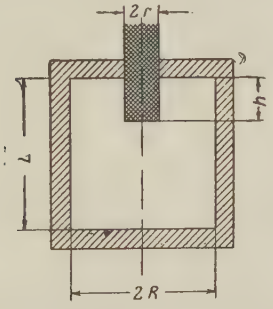


Figure 1. Cross section of a resonator with a cylindrical rod.

It follows from (1) that the frequency-detuning curves are nonlinear and can be only roughly approximated by linear segments. It is also easy to see that the steep linear portions are located near the points $h/\lambda_W=0.25;\ 0.75;\ \ldots$, while the gently sloping portions are near the points $h/\lambda_W=0.5,\ 1,\ \ldots$. The physical meaning of this is obvious. In the H_{01n} mode the only nonvanishing component on the resonator axis is the longitudinal one, H_z, which is distributed sinusoidally along the axis. When the end of the rod moves through the region of node of H_z $(h/\lambda_W=0.5,\ \ldots)$.

It can be shown that all these points will be points of inflection of the frequency-detuning curve. It can also be shown that if we approximate the steep portion of this curve about the point $h/\lambda_W=0.25$ by a line tangent to this curve at the point of inflection, the nonlinearity will become quite small if the width $\Delta h/\lambda_W$ of the portion is less than 0.2 (Figure 2), and it is therefore advantageous to use precisely this portion for retuning the resonator in the H_{01n} mode.

Differentiating (1) at the point $x=\pi$ $(h/\lambda_W=0.25)$, we obtain the equation for the tangent in this portion in the form

$$\Delta f=\frac{f_0}{L}G\Delta h, \qquad (2)$$

where

$$G=\frac{1}{J_0^2(\xi_{01})}\left(\frac{r}{R}\right)^2(\gamma\beta_z-\beta_1);$$

L is the resonant length of the unperturbed resonator; f_0 is the frequency of the resonator with the rod inserted a quarter wavelength.

Formula (2) enables us to determine the tuning band for specified rod dimensions, and conversely, the rod dimensions for a specified band. Formula (2) enables us also to calculate the dimensions of the rod for a given resolving power on the linear portion. Several values of G are listed in Table 1.

Table 1

r/R	λ w /r				
	0,9	1,0	1,1	1,2	1,3
0.05	0,005	0,006	0,007	0,008	0.009
0.10	0.017	0.021	0.026	0.031	0.037
0.15	0.033	0.042	0,052	0.063	0.075
0.20	0.048	0.063	0,080	0.098	0,017

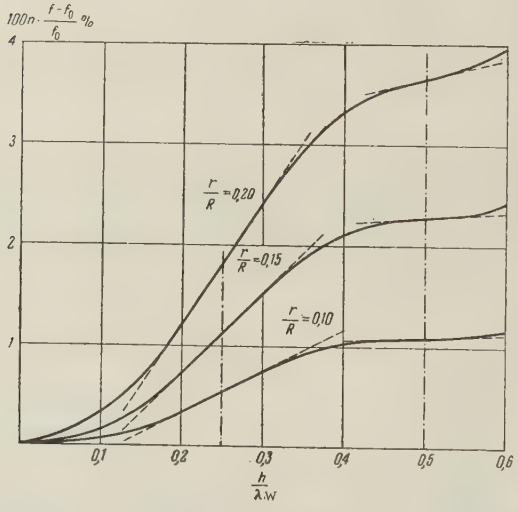


Figure 2. Frequency detuning of a resonator in the H_{01n} mode as a function of the relative insertion of the rod.

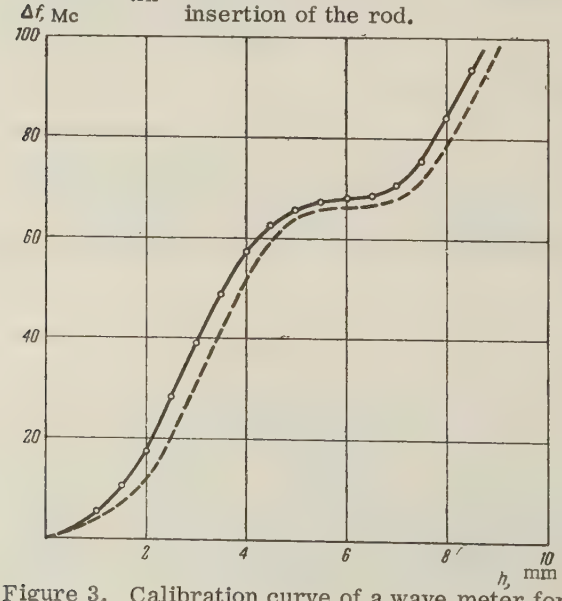


Figure 3. Calibration curve of a wave meter for the 8 mm band, r/R = 0. 08 [solid curve — experimental; dashed curve — calculated by formula (1)].

An experimental verification of the theoretical formulas is exceedingly simple. We calibrated the depth of insertion of the rod against the frequency and compared the calibration curve with the theoretical one. The results of such a comparison, made for a resonator (wave meter) for the 8 mm band, are shown in Figure 3. The construction of the wave meter has been described in [3]. It is seen from Figure 3 that the experimental and the theoretical curves are very close to each other, except that the experimental curve is shifted toward the smaller insertion depths. For this wave meter the ratio is $r/R = 0.08$. At greater values of r/R, the discrepancy between the experimental and theoretical curves becomes more appreciable, but the slope of the experimental curves in the linear portion (a quantity inverse to the resolving power, or the tuning rate) is in good agreement with the value of the tuning rate calculated by formula (2). The results of the comparison of the calculated and measured tuning rates of different investigated wave meters are summarized in Table 2.

In the first line of Table 2 are given the results for a wave meter for the 8 mm band [3]; the second line is for the 3 mm band. In both wave meters the unperturbed resonator was five half-wavelengths long ($n = 5$). The third line gives the results for a wave meter for the 1 cm band, described in [1]. The large tuning rate of this wave meter compared with the two others is due to the small value of $n\,(n = 1)$.

Table 2

r/R	$\Delta f/\Delta h$, Mc/mm		Error percent
	calculated	measured	
0.08	21.04	21.2	0.7
0.16	4.29	4.40	2.5
0.19	179	185 [1]	3.3

The data of Table 2 proved the correctness of formula (2) for the tuning rate on the linear portion. It can also be noted from Table 2 that the error in the calculation increases as the square of the ratio of the radii, $(r/R)^2$.

These results show the applicability of formula (1), which has been derived in a simple fashion, for an estimate of the frequency band of a wave meter with a rod, and the applicability of formula (2) for a reliable calculation of the resolving power on the linear portion of the frequency-detuning curve.

REFERENCES

1. Measurement Techniques at Centimeter Wavelengths, translated from the English, Soviet Radio Press, 1949, Part 1, page 120.
2. A.G. Gurevich, Hollow Resonators and Waveguides, Soviet Radio Press, 1952, page 98.
3. G.D. Burdun, E.B. Zal'tsman, V.E. Poyarkova, Apparatus for the Measurement of the Dielectric Constant and the Dielectric Loss Angle Tangent in the 8 mm Band, Collection "Hundred Years Since the Birth of A.S. Popov," Academy of Sciences Press, 1960.

Submitted to the Editors 28 October 1960

COUPLING COEFFICIENTS FOR A BENT COAXIAL WAVEGUIDE OF ROUND CROSS SECTION

O.Ye. Shushpanov

When a H_{01} wave passes through a bend in a round waveguide, part of the energy of this wave goes into energy of other modes. Energy is particularly transferred from the H_{01} mode into the E_{11} mode. To reduce the coupling between the H_{01} and the E_{11} modes at the bend of the waveguide, it is possible to introduce an internal impedance wire, which removes the degeneracy. The resultant coaxial system then supports a cable (transverse) electromagnetic mode and waveguide modes. We confine ourselves to the case when a perfect conductor is introduced into the bend of the waveguide. In article [1] a system of ordinary differential equations of first order was derived for the field at the bend of the waveguide. The coefficients of this system characterize the coupling between the wave incident on the bend and the

waves produced at the bend. These coefficients are called coupling coefficients. We derive below formulas for the coupling coefficients of the H_{01} mode and the transverse mode with other waves in a coaxial waveguide.

Let us examine the bend of a coaxial waveguide of round cross section with perfectly conducting walls. Let $\underline{a}$ be the radius of the external cyclinder, and b the radius of the internal cylinder. In such a waveguide, as in a round waveguide without an axial conductor, both E and H modes exist. We denote by E_{mn} the coaxial-waveguide E mode that goes as $b \to 0$ into the E_{mn} mode of a waveguide of radius $\underline{a}$ without an internal wire. We define the H_{mn} mode analogously.

The H_{01} mode is coupled only with the E_{11} and H_{1n} modes; the H_{01} and E_{11} modes remain degenerate in the coaxial system.

The coupling coefficient of the H_{01} and E_{11} modes, as in a waveguide without an axial conductor, has the form

$$F = \frac{ika}{\sqrt{2\mu}}. \tag{1}$$

For a coaxial waveguide μ is the first root of the equation

$$J_1(\mu)\, N_1\left(\frac{b}{a}\mu\right) - J_1\left(\frac{b}{a}\mu\right) N_1(\mu) = 0. \tag{2}$$

For a waveguide without an internal wire, μ is the first root of the equation $J_1(\mu) = 0$, that is, $\mu = 3.83$. Thus, for example, when $n = 0.5$ the coupling coefficient (1) of the H_{01} mode with the E_{11} mode for the bend of the coaxial waveguide is 1.7 times smaller than for the bend of a waveguide without an axial conductor.

The coupling coefficient of the H_{01} mode with the H_{1n} mode has the form

$$F = \frac{i(h_1 + h_n)^2 g_1 g_n^2}{\sqrt{2}(g_1^2 - g_n^2)^2 h_1} \frac{1}{\sqrt{N_1^2(g_1 b) - N_1^2(g_1 a)}} \times \\ \times \frac{aN_1(g_1 a)\, N_1'(g_n a) - bN_1(g_1 b)\, N_1'(g_n b)}{\sqrt{N_1'^2(g_n b)\, b^2 (g_n^2 a^2 - 1) - N_1'^2(g_n a)\, a^2 (g_n^2 b^2 - 1)}}, \tag{3}$$

where g_1, h_1, and g_n, h_n are respectively the wave numbers for the H_{01} and H_{1n} modes, with $g_1 = \mu_1/a$; μ_1 is the first root of Equation (2) and g_n is the n-th root of the equation

$$J_1'(g_k a)\, N_1'(g_k b) - J_1'(g_k b)\, N_1'(g_k a) = 0. \tag{4}$$

In the limiting case, when the radius of the internal wire tends to zero, we obtain from (2) the coupling coefficients for the H_{01} and H_{in} modes in a waveguide without an axial conductor [2].

The transverse (cable) wave at the bend is coupled only with the H_{1m} and E_{1n} modes. The coupling coefficients have the following respective forms

$$F = -\frac{ik(h_m + k)}{2h_m g_m^2} \frac{b^2 N_1'(g_m b) - a^2 N_1'(g_m a)}{ab\sqrt{\ln \frac{a}{b}}} \times \\ \times \frac{1}{\sqrt{N_1'^2(g_m b)\, b^2 (g_m^2 a^2 - 1) - N_1'^2(g_m a)\, a^2 (g_m^2 b^2 - 1)}}, \tag{5}$$

$$F = \frac{i2k(h_n + k)}{h_n g_n^2} \frac{aN_1(g_n a) - bN_1(g_n b)}{ab\sqrt{\ln \frac{a}{b}} \cdot \sqrt{N_1^2(g_n b) - N_1^2(g_n a)}}, \tag{6}$$

where g_m is the m-th root of (4); $g_n = \mu_n/a; \mu_n$ is the n-th root of (2). In the limiting case as $b \to 0$, expressions (5) and (6) tend to zero.

The author is grateful to B. Z. Katsenelenbaum for his interest in the work.

REFERENCES

1. B. Z. Katsenelenbaum, Bent Waveguides With Inhomogeneous Dielectrics, Radiotekhnika i Elektronika, 1958, 3, 5, 634.
2. N. P. Kerzhentseva, On the Propagation of Electromagnetic Waves in Bent Waveguides of Round Cross Section, Radiotekhnika i Elektronika, 1958, 3, 5, 649.

Submitted to the Editors 15 March 1961

SECONDARY AND PHOTOELECTRONIC EMISSION IN THE ADSORPTION OF BERYLLIUM AND SILVER ON BARIUM*

I. M. Bronshteyn, B. S. Frayman

1. It is known that when atoms of alkali and alkaline-earth metals are adsorbed on metallic targets, the work function φ of the latter decreases, and the coefficient of secondary electron emission σ increases. In particular, it was shown in [1] that upon adsorption of barium on molybdenum the maximum values of σ and of the photoelectronic emission depend on the degree of coating and are equal in magnitude.

Usually the reduction in the work function upon adsorption of atoms of electropositive metals on electronegative metals is attributed to the formation of an electric double layer, the positive side of which is on the outside. Consequently, the adsorption of atoms of electronegative metals on electropositive metals should lead to an increase in the work function, for in this case the double electric layer on the surface will be turned with its negative side to the outside. However the data given below, pertaining to measurements of the photoelectronic current and of σ in the adsorption of beryllium atoms and silver atoms on the surface of barium, show that this is not the case, and at the optimal degree of coating, the work function not only does not increase but, to the contrary, decreases.

2. The instrument, techniques and the procedure of the measurements are described in [2, 3]. The coefficient of secondary electron emission σ in the adsorption of barium on the surface of silver was measured at $5 \cdot 10^{-8}$ and $4 \cdot 10^{-9}$ mm Hg. The values of σ for barium and silver and variations with thickness $\sigma(d)$, were independent, within the limits of error, of the vacuum conditions [3]. The photocurrent was measured in the light from an incandescent tungsten helix.

Figure 1 shows $\sigma(t)$ and $I_{ph}(t)$ curves for the adsorption of barium on silver. Were σ_{Ba} to have the same value as σ_{Ag}, then the coefficient of secondary electron emission would decrease monotonically from σ_{Ag} to σ_{Ba} (dashed curve). In fact, when approximately half the monolayer of barium is adsorbed, the work function is already appreciably decreased, and σ increases approximately by 40 percent to a maximum. Upon further increase in the degree of coating of the surface with the adsorbed atoms of barium, σ decreases in spite of the reduction in φ, because the efficiencies of the forward and backward electron currents are lower in barium than in silver. This reduction in σ thus predominates over the increases in σ due to further reduction in φ.

Figure 2 shows curves of $\sigma(t)$ and $I_{ph}(t)$ for the adsorption of silver on barium. As can be seen in this case for a monolayer of silver on barium, there is also a noticeable decrease in the work function and an increase in the photocurrent by approximately 1.5 times. Because of this, $\sigma(t)$ likewise does not increase monotonically, but passes through a maximum. The maxima of the $\sigma(t)$ and $I_{ph}(t)$ curves are shifted appreciably because the efficiencies of the forward and reverse currents of the electrons are much greater in silver than in barium. Therefore, in spite of the sharp decrease in the photocurrent (increase in the work function), the increase in σ still continues.

* Reported at the Interdepartmental Seminar on Cathode Electronics, 6 February 1961.

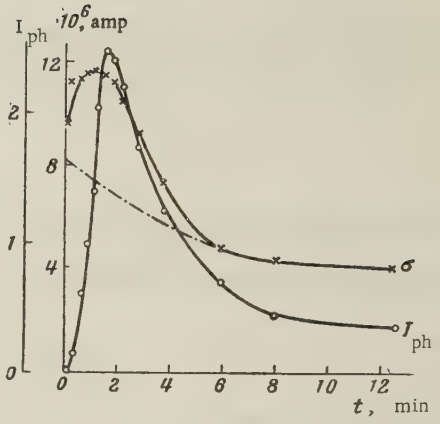


Figure 1. Curve showing the dependence of the photocurrent $I_{ph}(t)$ and the coefficient of secondary electron emission $\sigma(t)$ on the time of sputtering of the barium layer on the silver. The dashed curve (in this figure and in the other) shows the possible course of $\sigma(t)$ for equal work functions of the layer and the base. E_p=800 v.

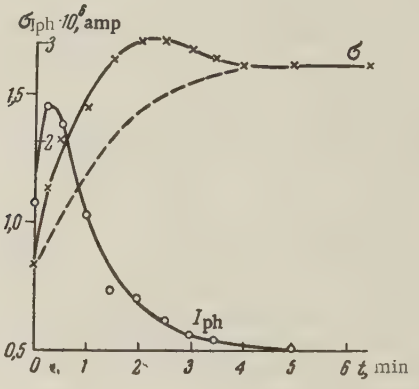


Figure 2. Curves showing the dependence of the photocurrent $I_{ph}(t)$ and the coefficient of secondary electron emission $\sigma(t)$ in the adsorption of silver atoms on barium. E_p = 800 v.

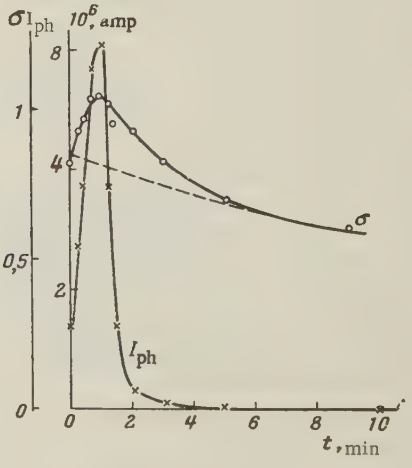


Figure 3. Curves showing the dependence of the photocurrent $I_{ph}(t)$ and of the coefficient of secondary electron emission $\sigma(t)$ in the adsorption of atoms of beryllium on barium. E_p=500 ev.

Figure 3 shows the curves of $\sigma(t)$ and $I_{ph}(t)$ for the adsorption of beryllium on barium. Because the efficiencies of the electron currents in beryllium and in barium do not differ greatly from each other in this case, the positions of the maxima of the photocurrent and of the coefficient of secondary electron emission σ coincide.

An analogous variation of $\sigma(t)$ is observed in the adsorption of barium on beryllium, of beryllium on calcium, of titanium on barium and calcium, and of barium and calcium on titanium.

3. From the foregoing results it is seen that the work function of the base can decrease not only upon adsorption of atoms from electropositive metals, but also upon adsorption of atoms of electronegative metals. It is obvious that this effect cannot be attributed to the formation of an electric double layer on the surface of the metal. It appears to us that the reduction in the work function upon adsorption of atoms of electronegative metals is due to the field of irregularities which exist on an irregular surface. In this case a change in the work function should manifest itself all the more strongly, the greater the difference in the work functions of the adsorbed metal and the base metal. In the case of beryllium, we have $_{Be}\varphi_{Be} \simeq 2\varphi_{Ba}$ [4] and I_{ph} increases by five times, whereas in the case of adsorption of silver it increases only by 1.5 times. As the dimensions of the homogeneous surface increase, the field of the irregularities over the surface becomes weaker. This is clearly seen in the adsorption of beryllium on barium; in this case the photocurrent decreases sharply even when the beryllium layer is one and a half atoms thick.

REFERENCES

1. J.H. deBoer, H. Bruining, Physica, 1939, 6, 941.
2. I.M. Bronshteyn, R.B. Segal, Fiz. Tverd. Tela. 1959, 1, 10, 1489.
3. I.M. Bronshteyn, B.S. Frayman, Doklady AN SSSR, 1960, 135, 1907.
4. A.P. Komar, V.P. Savchenko, V.N. Shrednik, Radiotekhnika i Elektronika, 1960, 5, 8, 1211.

Submitted to the Editors 4 February 1961

INVESTIGATION OF TRANSISTOR PULSE GENERATORS WITH DELAYED FEEDBACK

V.N. Yakovlev

It is known [1] that vacuum-tube regenerative oscillators make it possible to obtain pulses of nanosecond duration with a repetition frequency of some tens of megacycles. However, owing to the negligibly small utilization of the supply voltage, these generators are not widely used in practice.

Investigations showed that transistor oscillators with delayed feedback, unlike vacuum-tube oscillators, can be extensively used to obtain bell-shaped pulses of duration less than 0.1 μsec and with a highly stable repetition frequency of several megacycles. The pulse amplitude in this case can amount to 50 — 75 percent of the supply voltage.

The principle underlying the construction of transistor oscillators with delayed feedback does not differ essentially from the principle used for vacuum tube oscillators. These, as is well known [1, 2], consist of a strongly nonlinear noninverting broadband amplifier, a linear coupling two-port network, a delay line, and a network that ensures time regulation of the gain (Figure 1). The nonlinear amplifier can be a two-stage resistance-coupled amplifier, one stage of which can operate in the linear mode [1], and a single-stage transformer-coupled amplifier. The role of the coupling two-port is assumed by either an interstage RC coupling network or a transformer. The gain is usually regulated [1] with the aid of an RC amplifier bias network.

Depending on the choice of the amplifier operating conditions, on the delay-line parameters of the coupling two-port and the bias RC network, and finally on the form of the external disturbance signal, a great variety of steady-state oscillation waveforms can be obtained.

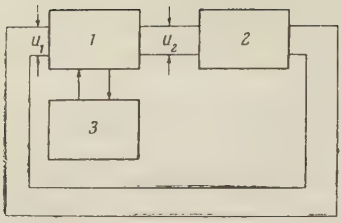


Figure 1. Block diagram of oscillator: 1) nonlinear amplifier; 2) coupling two-port and delay line; 3) gain control.

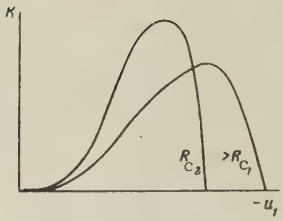


Figure 2. Plots showing dependence of the gain on the base bias voltage.

Transistor oscillators with delay have certain features due to the specific properties of the transistors. The dependence of the gain of an aperiodic p-n-p transistor amplifier on the bias voltage, has the form shown in Figure 2. At a low negative-bias voltage the transistor is at the cutoff threshold and the gain is close to zero. At a certain voltage it reaches a maximum. Further increase in the negative-bias voltage causes the collector current to reach values close to E_c/R_c, where E_c is the collector supply voltage and R_c the collector resistance. The voltage on the transistor collector then approaches zero, the operating point goes to the critical-mode line, and the gain drops to zero. Whereas in a vacuum-tube oscillator one can select the operating point for the production of bellshaped pulses only on the cutoff boundary [1], the operating point in a transistor oscillator can be chosen either on the cutoff boundary or on the transistor saturation boundary (Figure 2). Accordingly, the polarity of the pulses will be either positive or negative. An operating mode corresponding to the start of the transistor saturation can be readily realized, and should be given preference in the generation of bell-shaped pulses. The reason is that in this case the dependence $K = f(u_1)$ is more clearly pronounced, and consequently the pulse duration will be shorter.

The advantage of transistor oscillators with delayed feedback over relaxation transistor oscillators of the ordinary type (multivibrators and blocking oscillators) lies in the fact that deep saturation of the transistor is eliminated in the former. As regards the signal delay time, which takes place when the signal is amplified by a transistor amplifier, it adds to the delay-line time and exerts no influence whatever on the slope of the pulse fronts. It therefore becomes possible to realize the superior frequency properties of drift transistors and to obtain pulses with durations of about 0.05 microsecond. The maximum pulse repetition rate can be on the order of $5 \cdot 10^6$ pulses per second.

A regenerative transistor pulse oscillator with delay can be built with one (Figure 3b) or two transistors (Figure 3a). The voltage dividers of the two-stage oscillator, made up of resistors R_1, R_2, R_3, R_4 make it possible to select the operating mode for the amplifier stages. Inasmuch as the pulse polarities at the divider inputs are opposite, their operating modes are best made different. Thus, if negative-polarity pulses are applied to the collector of transistor T_2 (Figure 3a), the transistor operating condition must be close to saturation. Transistor T_1 should then be at the cutoff boundary. In this case it is possible to make use of the sharpening action of the two amplifier stages and obtain pulses of shorter duration. The emitter networks R_eR_c (Figure 3a) result in soft self-oscillation, stabilize the transistor operating conditions, and ensure regulation of the amplifier gain.

The oscillator shown in Figure 3b is simpler to adjust and more reliable in operation. The transformer is used to invert the phase of the transformed voltage. As in the preceding oscillator, the divider R_1, R_2 and the R_eC_e network ensure the choice of the required operating mode and the regulation of the gain. The resistance and the capacitor of the automatic bias network R_eC_e can be transferred to the base circuit. In this case the pulse amplititude increases and can reach 50 — 75 percent of the collector supply voltage. If a transformer with a third load winding is used it is possible to obtain the required amplitudes and polarities of the pulses.

The purpose of the experiment described was a qualitative check on the operating ability and effectiveness of these semiconductor oscillators with delayed feedback.

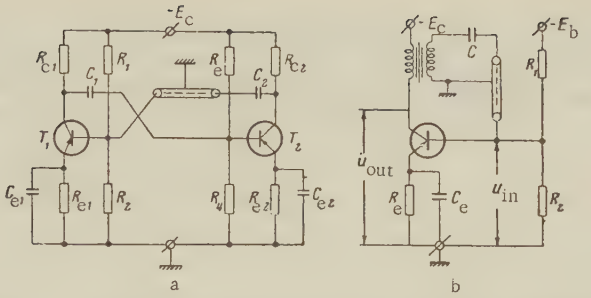


Figure 3. Principal diagram of two-stage (a) and single-stage (b) pulse oscillators with delayed feedback

The two-stage oscillator shown in Figure 3a uses type P-402 transistors with $E_k = -10$ v, $R_1 = 22$ kΩ, $R_2 = R_{c1} = R_{c2} = 1.2$ kΩ, $R_3 = 13$ kΩ, $R_4 = 2.2$ kΩ, $R_{e1} = 680$ ohms, $R_{e2} = 120$ ohms, $C_1 = C_2 = 6800$ μμf, $C_{e1} = 1000$ μμf, and $C_{e2} = 400$ μμf. If the delay line used is a high-frequency coaxial cable, it is possible to obtain on the collector of transistor T_2 pulses of almost bell-shaped form with negative polarity and minimum duration $t_{p2} = 0.05$ μsec and maximum repetition frequency of about 5 Mc. When the cable is replaced by a low-pass filter with lumped parameters, the pulse duration is increased somewhat. In this case the pulse amplitude was 1 — 1.5 v. The oscillator is quite sensitive to the parameters of the emitter networks R_eC_e. When their values are changed from those indicated above, new oscillation modes appear.

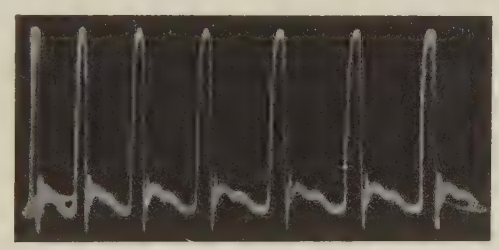


Figure 4. Pulse oscillograms

The single-stage oscillator shown in Figure 3b is more reliable and more efficient. It employs a P-402 transistor with $E_c = -10$ v, $R_e = 120$ ohms, $C_e = 510$ μμf, $R_1 = 2$ kΩ, and $R_2 = 1$ kΩ. It produced nearly bell-shaped pulses with an amplitude of about 3 — 4 v, duration about 0.08 μsec, and a maximum repetition frequency of about 4 Mc. The number of turns in the collector winding of the transformer is 45; the base winding has 15 turns; the core is a ferrite ring with an outside diameter of about 8 mm. Replacement of the cable by a low-pass filter, as in the first circuit, leads to a certain increase in pulse duration. The use of a P13a transistor in the single-stage oscillator yielded results close to those described above. This gives grounds for hoping that the single-stage oscillator with a low-pass filter can be widely used as an oscillator for short-duration pulses with a highly stable and high repetition frequency.

Figure 4 shows oscillograms of pulses of negative polarity, obtained with the aid of the two-stage oscillator (Figure 3a) using a delay line in the form of a low-pass filter with a pass-band $\Delta f = 5$ Mc, characteristic impedance $k\Omega\rho = 1$ kΩ, and delay time $\tau = 1$ μsec. The pulse repetition period on this oscillogram is 1 μsec.

The pulse duration increases slightly with increasing delay time, owing to the greater attenuation produced by the losses in the high-frequency components of the pulse spectrum. A change in temperature over a wide range produces only an insignificant change in the pulse waveform and exerts no influence on its repetition period.

The experimental tests have also shown that generation of bell-shaped pulses is also possible when an oscillator with negative delayed feedback is used. In this case the pulse duration depends essentially on the delay time. However, a detailed analysis of such an oscillator is beyond the scope of the present article.

REFERENCES

1. C.C. Cutler, Regenerative pulse generator, Proc, I.R.E., 1955, 43, 2, 140.
2. G.V. Glebovich, L.A. Morugin, Shaping of Nanosecond Pulses, Soviet Radio Press, 1958.

Submitted to the Editors 21 March 1960
after revision 29 May 1961

THERMIONIC PROPERTIES OF BARIUM HAFNATES AND RHENATES

B.V. Bondarenko, S.V. Yermakov, B.M. Tsarev

In connection with the results of our investigation [1] of the thermionic properties of barium tantalates, which were found to be somewhat better than those of barium tungstates, the question arose of the need for investigating the emission from barium rhenates and hafnates, too. We investigated hafnates of barium of the type $(BaO)_n (HfO_2)_m$ with composition $n : m = 2 : 1; 3 : 1; 5 : 1; 7 : 1$ and rhenates of barium $(BaO)_n (Re_2O_7)_m$ with composition $n : m = 1 : 1; 2 : 1; 3 : 1; 5 : 1; 7 : 1$.

An investigation of barium rhenates and hafnates with different contents of barium oxide was essential to elucidate the effect of the barium oxide content on the thermionic properties of the double oxide, and also for the determination of the compositions of the basic types of barium rhenate and hafnate, which are stable in vacuum at the working temperatures.

The technology of preparation of the hafnates and rhenates, and also the procedure used to measure their thermionic properties, were similar to those already described [1] for the tungstates and tantalates of alkaline-earth metals. The backing for the cathodes was a tungsten strip, first cleaned by heating in vacuum. The temperature of the core was determined with the aid of a tungsten-iridium thermocouple.

The cathode activation process consisted of a prolonged heating while drawing the emission current, starting with the temperature at which weak emission, amounting to 10^{-8}—10^{-7} amp/cm^2 appears, and ending with a temperature above which the irreversible decrease in emission due to the increase in work function sets in. Up to this value of the temperature, which was designated as the activation temperature, there was as a rule a sharp decrease in the work function during the activation process. By way of an example, we show the activation curves for the basic barium rhenate and hafnate (see Curves a in Figures 1 and 2).

After the end of the cathode activation, we measured the emission in a wide range of temperatures. Several runs up and down the temperature scale were taken in this case until stable and reproducible emission currents were obtained for each value of the temperature (the stable-emission state curves, marked b, are also shown in Figures 1 and 2).

Table

Thermionic properties of barium hafnates and rhenates

Composition of compound	Activation temperature °K	φ min at T °K	φ_T, ev, at $T = 1250$ °K	φ_0, ev	A, amp/cm^2 · deg^2
$2BaO \cdot HfO_2$	1200	2.12 1000	2.20	1.57	$7.5 \cdot 10^{-1}$
$3BaO \cdot HfO_2$	1200	2.08 1000	2.18	1.49	$5.4 \cdot 10^{-1}$
$5BaO \cdot HfO_2$	1200	1.98 950	2.17	1.42	$1.15 \cdot 10^{-1}$
$7BaO \cdot HfO_2$	1200	1.96 900	2.16	1.35	$6.4 \cdot 10^{-2}$
$BaO \cdot Re_2O_7$	1450	2.92 1100	2.94	1.69	$1.1 \cdot 10^{-3}$
$2BaO \cdot Re_2O_7$	1300	2.55 1000	2.68	1.43	$1.1 \cdot 10^{-3}$
$3BaO \cdot Re_2O_7$	1300	2.47 1000	2.57	1.57	$1.1 \cdot 10^{-1}$
$5BaO \cdot Re_2O_7$	1350	2.35 1000	2.44	1.67	$1.15 \cdot 10^{-3}$
$7BaO \cdot Re_2O_7$	1350	2.27 1000	2.38	1.51	$7.4 \cdot 10^{-2}$

All the substances investigated by us had a minimum work function in accordance with their semiconductor properties, described by the simplest model of a donor semiconductor.

The linear temperature dependence of the work function in the region above the temperature at which the minimum work function is observed, has enabled us to calculate the values of the constants φ_0 and A of the line equation of Richardson.

$$j = AT^2 \exp\left[-\frac{e\varphi_0}{kT}\right]$$

From the values of $\Delta\varphi/\Delta T$, using the equations

$$A = A_0 \exp\left[-\frac{e}{k}\frac{\Delta\varphi}{\Delta T}\right],$$

$$\varphi_0 = \varphi_T - \frac{\Delta\varphi}{\Delta T}T,$$

where A_0 is a universal constant with value $4\pi mk^2e/h^2 = 120.4$ amp/cm^{2}deg^2, and φT is the work function in the following emission equation:

$$j = A_0 T^2 \exp\left[-\frac{e\varphi_T}{kT}\right].$$

The thermionic properties of the barium hafnate and rhenate compounds investigated by us are listed in the Table. Comparing the values of the work function φT at 1200°K for barium hafnate $2BaO \cdot HfO_2$ ($\varphi_T = 2.20$ ev), tantalate $5BaO \cdot Ta_2O_5$ ($\varphi_T = 2.30$ ev), tungstate $3BaO \cdot WO_3$ ($\varphi_T = 2{,}61$ ev), and rhenate $7BaO \cdot Re_2O_7$ ($\varphi_T = 2.38$ ev) we see that the rhenates and especially the hafnates of barium have several better emission properties than the corresponding tungstates and even tantalates. However, for final recommendations with respect to the possibility of using hafnates and rhenates of barium as active substances of various types of pressed thermionic cathodes, additional research must be carried out on their volatility and thermal stability under working conditions.

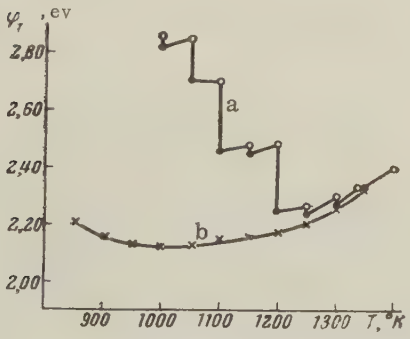


Figure 1. Activation (a) and stable-emission (b) curves for the basic composition of barium hafnate $2BaO \cdot HfO_2$.

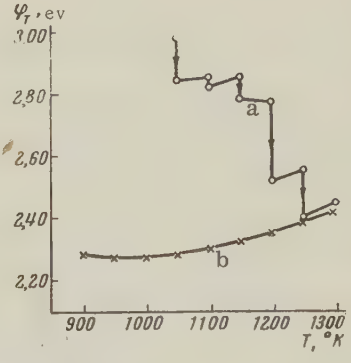


Figure 2. Activation (a) and stable-emission (b) curves for the basic composition of barium rhenate $7BaO \cdot Re_2O_7$

REFERENCES

1. B.V. Bondarenko, E.P. Ostapchenko, B.M. Tsarev, Thermionic Properties of Tungstates of Alkaline Earth Metals, Radiotekhnika i Elektronika, 1960, 5, 8, 1246.

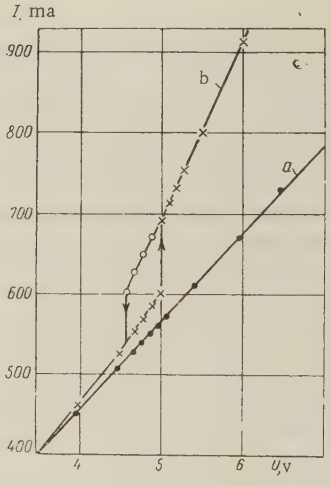
Submitted to the Editors 15 June 1960

ANOMALOUS VOLT-AMPERE CHARACTERISTIC OF ALLOYED GERMANIUM

F.J. Nagy

We observed in a specimen of germanium alloyed with gold and having fused indium contacts an anomalous volt-ampere characteristic, and also generation of sinusoidal oscillations in a network consisting of the specimen, a load resistance, and a battery.

The single-crystal homogeneous p-type germanium specimen contained $3 \cdot 10^{15}$ cm^{-3} of gold impurities at a degree of compensation of 0.01. The specimen measured $1 \cdot 1.5 \cdot 12$ mm and was placed in a cryostat which reduced the temperature to 80°K. At this temperature the specimen had a resistance ~ 9 kilohm. The form of the volt-ampere characteristic depended on the polarity of the voltage applied to the specimen. At one polarity the characteristic is linear and has no singularity (see the figure, Curve a). Reversal of the polarity of the voltage applied to the specimen causes an appreciable change in the form of the volt-ampere characteristic, viz., an increase in the voltage on the specimen produces a sudden, sharp increase in the current (Curve b) at a certain point, corresponding to 4.98 v. Upon further increases in the voltage the current continues to increase linearly, but with a different (greater) slope of the characteristic.



When the characteristic is plotted in the reverse direction, the jump in current occurs at a smaller voltage, -4.6 v. The volt-ampere curve forms a characteristic hysteresis loop.* By connecting the specimen in series with a resistive load ($R_{load} \simeq R_{specimen}$) and a battery (with the total wiring capacitance amounting to about 100 $\mu\mu$f) and by regulating the voltage of the battery it was possible to observe the occurrence of the stable generation of sinusoidal oscillations with an approximate frequency of 500 kc and an amplitude ~ 70 mv. Oscillations set in only at the voltage polarity corresponding to the presence of hysteresis.

When specimen was illuminated with white light (from an incandescent lamp) through a window in the cryostat, the oscillations ceased. After the light was turned off the oscillations ceased. After the light was turned off the oscillations resumed, the amplitude of the oscillating voltage building up slowly from zero to its initial value with a time constant ~ 1 minute. The fact that the phenomenon described disappeared when the polarity of the voltage was reversed indicates that it is apparently connected with processes that occur on the contacts.

REFERENCES

1. R. Bube, J. Appl. Phys., 1960, 31, 12, 2239.

Submitted to the Editors 8 April 1961

* Analogous volt-ampere characteristics were observed by Bube [1] in polycrystalline CdS specimens.

At the request of the National Committee of the Soviet Union on Automatic Control, the Editor offers the following information.

In March 1961 there there was held in Bergen (Norway) a session of the Executive Council of the International Federation on Automatic Control (IFAC), under the chairmanship of the IFAC president, Prof. A. M. Letov. The Executive Council of the International Federation on Automatic Control (IFAC), under the chairmanship of the IFAC president, Prof. A. M. Letov. The Executive Council of IFAC adopted the following resultion concerning the holding of the second IFAC congress.

The second IFAC congress will be held in Basel (Switzerland) in September 1963.

A. The program will cover the following scientific disciplines:

1. Theory of automatic control: a) discrete systems; b) stochastic processes; c) optimal systems; d) self-adaptive systems; e) reliability theory.

2. Applications of automatic control: a) investigations of the dynamics of processes; b) study of problems in the automatization of industry with the aid of digital and analog devices that either participate or do not participate in the process; c) use of optimizing and self-adaptive control systems.

3. Elements: a) new and efficient devices; b) estimate of element reliability.

4. Other topics: a) education; b) terminology; c) bibliography

This program does not exclude other topics, but not less than 80 percent of all papers should pertain to theory and applications, and not more then 20 percent to elements and other subjects.

B. The total number of approved papers must not exceed 100. Each paper should contain not more than 30,000 printed symbols, including a resume in two or three languages and illustrations.

C. The selection of papers will be made by the National Committees of IFAC, who should: a) invite specialists to present papers on the aforementioned topics; b) find specialists capable of expressing opinions and preselecting the papers; c) advise the authors that the papers should be sent to the Soviet Union National Committee on Automatic Control (Moscow, I-53, Kalanchevskaya 15-a) not later than 31 December 1961. The author can present papers in Russian, English, French, or German, each paper being accompanied by a summary not longer than 200 words both in the original language and in Russian or English.

The final selection of papers for the Second IFAC Congress will be made by the Committee on Papers of the IFAC, comprising Prof. Ed. Gehreke (Switzerland), chairman, and the chairmen of the Technical Committees of IFAC as members. The Committee on Papers of the IFAC will be guided by the following principle in its decisions: 1) timeliness of topic; 2) importance and novelty of the results obtained; 3) clarity of exposition.

Chairman, USSR Nation Committee on Automatic Control — Acad. V. A. Trapeznikov

---

RADIO ENGINEERING AND ELECTRONIC PHYSICS

The English Edition of Radio Engineering and Electronic Physics is mailed to subscribers within 18 weeks after the publication of the original Russian issue.

RADIO ENGINEERING AND ELECTRONIC PHYSICS

РАДИОТЕХНИКА И ЭЛЕКТРОНИКА

NUMBER 10 1961

CONTENTS